Technical Report DSL Forum TR-059

DSL Evolution - Architecture Requirements for the Support of QoS-Enabled IP Services

September 2003

Produced by:
Architecture & Transport Working Group
Editor: Tom Anschutz, BellSouth Telecommunications

Working Group Co-Chair: David Allan, Nortel Networks
Working Group Co-Chair: David Thorne, BT

Abstract:
This Working Text will outline an evolution of mass market DSL services to deliver multiple levels of QoS enabled IP layer services to DSL subscribers. In support of this service evolution, a reference architecture and supporting requirements are included that outline the interface specifications needed from a subscriber or a Service Provider to access these new services.

Revision History	Date	Reason for Update
Version 1	September, 2002	Created new document based on technical requirements sections of dsl2002.213
Version 2	October, 2002	Comments and minor updates from Oct. 8 call
Version 3	December, 2002	Comments and minor updates from Nov. 12 call
Version 4	December, 2002	Comments and updates from December DSLF meeting in San Francisco – except, this version does not yet number requirements – as was agreed.
Version 5	February, 2003	Introduced numbered requirements, removed "RSVP-like," accepted reviewed changes from V4, and incorporated contributions from January interim teleconference.
Version 6	March, 2003	Moved Phase-2 QoS and Multicast information into Appendix as agreed in February DSLF meeting in Dallas.
Version 7	March, 2003	Collected Interim Meeting Comments (largely editorial) in preparation for straw ballot.
Version 8	May, 2003	This version captures straw ballot comments resolved in Lisbon and is intended to be the text to work from on the Straw ballot comments finalization call.
Version 9	June, 2003	This version captures straw ballot comments resolved in interim meeting and is intended to be the text for letter ballot.

Table of Contents

1 PURPOSE AND SCOPE 1

1.1 PURPOSE 1
1.2 SCOPE 1
1.3 REQUIREMENTS 2

2 PRODUCTS AND SERVICES 2

2.1 SERVICE GOALS 2
2.2 PRODUCT AND SERVICE LISTS 2

3 FUNCTIONAL ASSUMPTIONS 4

3.1 KEY TERMINOLOGY 4
3.2 BROADBAND PROVIDER REFERENCE DEFINITIONS 6
3.3 INTERFACES 8
3.3.1 A10-ASP Interface 9
3.3.2 A10-NSP Interface 9
3.3.3 U Interface 9
3.3.4 T Interface 9

4 REFERENCE ARCHITECTURE 9

4.1 LOGICAL REFERENCE ARCHITECTURE 9
4.2 LOGICAL ELEMENTS AND INTERFACES 11
4.2.1 Application Service Provider Network 11
4.2.2 A10-ASP Interface 12
4.2.3 Network Service Provider Network 14
4.2.4 A10-NSP interface 14
4.2.5 Regional/Access Network 18
4.2.6 U Interface 22
4.2.7 Customer Premises Network 24
4.2.8 T Interface 26

5 QUALITY OF SERVICE 28

5.1 INTRODUCTION 28
5.1.1 Goals 29
5.1.2 Assumptions 29
5.2 TRAFFIC ENGINEERING OF BEST EFFORT SERVICE 29
5.2.1 Theory of Operation 29
5.3 QOS ARCHITECTURE - A TWO-PHASED APPROACH 30
5.3.1 Phase 1 QoS Mechanisms 30
5.3.2 Phase 2 QoS Mechanisms 33

6 SERVICE LEVEL MANAGEMENT 36

6.1 INTRODUCTION 36
6.2 NETWORK PERFORMANCE METRICS 36
6.3 OPERATIONAL METRICS 36

7 SERVICE MANAGEMENT 36

7.1 SUBSCRIBERS 37
7.2 SERVICE PROVIDERS 37

GLOSSARY 38

APPENDIX A REFERENCES 41

APPENDIX B INFORMATIVE EXAMPLE OF QUEUING ARCHITECTURES FOR RG AND BRAS 42

B.1 EXAMPLE QUEUING ARCHITECTURE FOR RG 42
B.2 EXAMPLE QUEUING ARCHITECTURE FOR A BRAS THAT CAN ALSO SWITCH ATM 43

APPENDIX C INFORMATIVE APPENDIX ON SIGNALED QOS 47

C.1 SIGNALED QOS MECHANISMS 47
C.1.1 Signaled QoS Assumptions 47
C.1.2 Diffserv Assumptions 48
C.1.3 Traffic Engineering Requirements 48
C.1.4 Admission Control 48

Table of Figures

Figure 1 – DSL Network Components 7
Figure 2 – Many-to-Many Access 7
Figure 3 - ATM based Regional and Access Network Providers 10
Figure 4 - IP Enabled Regional Network 11
Figure 5 - A10-ASP Interface 12
Figure 6 - ASP Protocol Stack with QoS 13
Figure 7 - A10-NSP Interface Supporting L2TP Connection 15
Figure 8 - L2TP Protocol Stack 15
Figure 9 - A10-NSP Interface Supporting IP Routed Connection 16
Figure 10 – Routed IP Protocol Stack with QoS 17
Figure 11 - Components of the Regional/Access Network 18
Figure 12 - Aggregation function of Regional Network 19
Figure 13 – Access Node Architecture Variations 22
Figure 14 – U Interface 23
Figure 15 – U Interface Protocol Stack 24
Figure 16 - T Interface 27
Figure 17 - IP over Ethernet 28
Figure 18 - IP over PPP over Ethernet 28
Figure 19 – Best Effort TE 30
Figure 20 – QoS-enabled Network Topology 32
Figure 21 – Phase 2 with Policy-based profiles 34
Figure 22 – Queuing and Scheduling Example for RG 43
Figure 23 – Reference Topology for Queuing and Scheduling Example for a BRAS that can also switch ATM 44
Figure 24 – Queuing and Scheduling Example for a BRAS that can also switch ATM 46

1 PURPOSE AND SCOPE

1.1 Purpose

ADSL service providers are highly interested in advancing DSL to be the preferred broadband access technology by growing their networks, increasing the value provided by those networks, and expanding the market they can address. To do this they must address several critical needs, particularly:

- The service must become more accessible to end-users and to wholesale and retail partners.
- The service must address a wider market with:
  - Variable speeds,
  - Variable precedence arrangements – allowing some application's traffic to take precedence over others.
  - Specific support for IP applications (e.g. IP-QoS and multicasting),
  - Support for new business models that can include more types of service providers, and
  - Support for these new service parameters across multiple connections to different service providers from a single DSL subscriber.
- The service must be competitive with alternative access technologies such as cable modem.

While adopting new architectures, like FSVDSL, may also fulfill these needs, perhaps even better than the architecture defined here, it is also important to realize that much ADSL has already been deployed, and that the current business imperatives may cause ADSL service providers to try to make more of what they already have than to try massive upgrades along with the massive capital investment they usually bring.

Therefore, there is also a critical need to provide a standard evolution path for the embedded base of ADSL.

The purpose of this work and the new service models is to provide a more common architecture and set of service interfaces to address these critical needs. Adhering to this architecture and to the services and service models set forth both here and in WT80 simplifies and unifies the way for all types of service providers to obtain ADSL end-user customers whether they sell access to networks, applications, or content.

The anticipated outcome for employing this specification, as well as others that build from it, is that it will:

- Reduce the number of alternative interfaces to ISPs/ASP and end users, in order to reduce costs through common interconnection.
- Establish guidelines for developers and vendors, so they can build equipment that support common service requirements.
- Improve the ability to introduce end-to-end services and applications worldwide, so that similar services can interwork across various service providers' networks.

1.2 Scope

This document presents an architecture for evolving DSL deployment and interconnection including the LAA and PTA architectures defined in TR-25. It outlines a common methodology for delivering QoS-enabled applications to DSL subscribers from one or more Service Providers. The business framework and drivers justifying this architectural evolution are described, in part, in WT-080. In the largest sense, the scope of this architecture is to provide IP-QoS and more flexible service arrangements to millions of users and thousands of service providers. And to do this to a useful extent, while pursuing only economic enhancements to existing ADSL networks.

While ADSL is useful for mass markets, segments and niches – this architecture addresses the mass market specifically. The approach, service models, and architecture are intended to scale to thousands of service providers, and many millions of end-users. The architecture does not detail approaches and techniques that might be appropriate to segments and niches, but does recognize that they might also be used in concert with

this approach. Similarly, local regulations, e.g. wiretapping, might apply to this and any architecture, but are beyond the scope of this document.

Many of the requirements levied on network elements and management systems are collected in this architecture, but they should not be taken as an exhaustive list of requirements for such elements. It is expected that other documents and standards will come forward to collect the requirements here, as well as those from other markets, segments, and niches in order to provide complete requirements for elements and systems that wish to be suitable in the DSL industry.

This architecture provides a high-level, evolving view of ADSL access. Because of this it provides more details about things that will happen sooner and fewer details about things that are several years and phases from fruition. Also, unlike a design, this architecture does not provide exhaustive instructions on how to develop and deploy networks or elements that adhere to the architecture. In fact, it identifies the need to develop and standardize new functions, features, and protocols in many later-phase areas.

1.3 Requirements

In this document, several words are used to signify the requirements of the specification. These words are often capitalized.

MUST	This word, or the adjective "REQUIRED", means that the definition is an absolute requirement of the specification
MUST NOT	This phrase means that the definition is an absolute prohibition of the specification.
SHOULD	This word, or the adjective "RECOMMENDED", means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications must be understood and carefully weighted before choosing a different course.
MAY	This word, or the adjective "OPTIONAL", means that this item is one of an allowed set of alternatives. An implementation that does not include this option MUST be prepared to inter-operate with another implementation that does include the option.

2 PRODUCTS AND SERVICES

2.1 Service Goals

Despite efforts to unify the architecture of Service Provider connections and to provide common service tiers, there has not been general support for a unified architecture. This proposal intends to increase the interest in such an architecture by increasing the number of service parameters available as well as by making those parameters more dynamic. Aside from variable dynamic bandwidth, this new architecture includes Quality of Service (QoS) and multi-application/multi-destination selection.
Service Providers benefit in that they will only need to develop one set of system interfaces for any carrier that adopts this architecture. By subscribing to these interfaces, Service Providers will now be able to develop applications that can take advantage of variable bandwidth and differentiated data traffic delivery that supports better than *best effort* traffic classes. Subscribers will be able to realize greater potential of their broadband data connections. This means that a subscriber can still use their Internet access as it exists today; yet additional bandwidth on their DSL line can be used to deliver other applications, such as direct corporate access, video chat and video conferencing, and various content on demand - be it movies, games, software, or time-shifted television programs. Finally, these applications can be given QoS treatment, so that business access, online gaming, and casual Internet access all share bandwidth appropriately. Both subscribers and Service Providers will be able to choose who provides the best service for a specific application, and what applications add the most value.

2.2 Product and Service Lists

This document presents a proposal for evolving DSL deployment and interconnection. It will outline a common methodology for delivering QoS-enabled applications to DSL subscribers from multiple Service Providers. These products and services are intended to address the mass market, and do not preclude additional niche or custom services that could be provided using the same infrastructure. Many of the current products offered

today either can be adapted to contain or already do contain the necessary software needed to support the proposed architectures contained within this document.
Also provided is a set of architectural requirements to support the proposed new service models. Some of the highlights include:

- IP-based services and QoS
- A single network control plane
- The migration of DSL regional transport to leverage newer, alternative technologies

The prevalent existing service model, where subscriber connections are delivered in a best effort fashion over end-to-end ATM PVCs, will continue to exist. However, this service model cannot support many of the improvements and benefits desired, including IP QoS, bandwidth on demand, and utilization of newer, alternative transport options.
This architecture supports the following service provider interconnection models, which are described WT-080:

- Subscriber access using PPPoE aggregated into L2TP tunnels delivered to Network Service Providers.
- Subscriber access using PPPoE or IP over Ethernet aggregated into VPNs delivered to Network Service Providers.
- Subscriber access using PPPoE or IP over Ethernet aggregated into a common, public, QoS-enabled IP network and delivered to Application Service Providers.

The DSL architecture and requirements put forward by this document enable the following product and service enhancements, which are described in WT-080.

- Bandwidth on Demand
- QoS, including QoS on Demand
- Many-to-Many Access
- Content Distribution

Network Service Providers will be able to benefit from the aggregation capabilities of the new DSL Access Networks described in this document. Specifically, this architecture will also permit:

- Traffic Aggregation: The end-to-end ATM PVC models, whether VPC or VCC, do not provide a scalable solution. L2TP and IP are used to provide better scalability and efficiency.
- Improved Transport: Currently most DSL transport is done over ATM connections. By offering other transport options, like Packet over SONET (POS) and Metro Ethernet, this architecture can provide better scalability, reduced overhead, and increased flexibility.
- Simpler Provisioning: Because they are not directly linked to provisioning transport, L2TP and IP delivery models can reduce the level of per subscriber provisioning.
- Differentiated Services: Up until now, almost all DSL transport has been best effort delivery. This new IP based architecture will permit Service Providers to offer differentiated treatment for certain traffic.
- Bandwidth Services: Up until now, most DSL access has been at a fixed rate that was selected at the time an access was provisioned. This architecture provides mechanisms that allow rates to be selected or changed more often and potentially on-the-fly.

- Increased Access: In previous architectures, Service Providers could only reach those subscribers with whom they had a direct relationship. These new architectures permit a subscriber to connect simultaneously to multiple Service Providers for a variety of services. Service Providers no longer need to be the sole provider to their subscribers.

- Standard Connections: Up until now, each access provider has had their own set of interfaces for Service Providers. This proposal defines common interfaces for NSPs and ASPs. This means that the Service Provider need only develop a single interface to get all of these features for many access providers. Also, subscriber connections will be similar among Access Providers, allowing common CPE to be more widely deployed.

Support for these new services will require a new set of network management interfaces. Both Service Providers and Subscribers will use these interfaces. Service Providers will be able to examine the network and see how their subscribers are provisioned. NSPs will also be provided an interface to control and troubleshoot subscriber connections.
Subscribers will be provided mechanisms for requesting these new services and indicating specific needs. These requirements will support applications and services like:

- Multicast audio and video media applications
- Video on demand applications
- Voice services
- Interactive gaming
- Variable bandwidth, both on demand ("Turbo" button) and by application

3 FUNCTIONAL ASSUMPTIONS

3.1 Key Terminology

The following definitions apply for the purposes of this document:

Access Network The Access Network encompasses the elements of the DSL network from the NID at the customer premises to the BRAS. This network typically includes one or more types of Access Node and often an ATM switching function to aggregate them.

Access Node The Access Node contains the ATU-C, which terminates the DSL signal, and physically can be a DSLAM, Next Generation DLC (NG-DLC), or a Remote Access Multiplexer (RAM). A DSLAM hub can be used in a central office to aggregate traffic from multiple remote physical devices, and is considered logically to be a part of the Access Node. When the term "DSLAM" is used in this document, it is intended to very specifically refer to a DSLAM, and not the more generic Access Node. The Access Node provides aggregation capabilities between the Access Network and the Regional Network. It is the first point in the network where traffic on multiple DSL lines will be aggregated onto a single network.

Behavior The externally observable characteristic applied to a traffic stream by a network element or system, for example assuring a minimum rate for a video stream or PPP session.

Broadband Remote Access Server (BRAS) The BRAS is the aggregation point for the subscriber traffic. It provides aggregation capabilities (e.g. IP, PPP, ATM) between the Regional/Access Network and the NSP or ASP. Beyond

	aggregation, it is also the injection point for policy management and IP QoS in the Regional/Access Networks.
Core Network	The center core of the Regional Network. The functions contained herein are primarily transport oriented with associated switching or routing capabilities enabling the proper distribution of the data traffic.
Downstream	The direction of transmission from the Access Node to the DSL modem.
Dropping	The process of discarding packets/cells based on specified rules, which may be the result of for example, a policing action or policy decision.
Edge Network	The edge of the Regional Network. The Edge Network provides access to various layer 2 services and connects to the Regional Network core enabling the distribution of the data traffic between various edge devices.
Layer 2 Tunnel Switch (L2TS)	The L2TS provides a second layer of PPP aggregation beyond the L2TP Access Concentrator (LAC). PPP sessions are switched between L2TP tunnels and are further aggregated and delivered to the NSP.
Loop	A metallic pair of wires running from the customer's premises to the Access Node.
Many-to-Many Access Sessions	The ability for multiple individual users or subscribers, within a single premises, to simultaneously connect to multiple NSPs and ASPs.
Microflow	A single instance of an application-to-application flow of packets, which may for example be classified by source address, source port, destination address, destination port and protocol id, or stateful means.
PVC Bundle	Two or more ATM PVCs (called a "bundle") are co-terminated on router endpoints. Each bundle co-termination is bound to a single IP interface. That is, the two (or more) PVCs appear to be a single "link layer" to the IP layer and so share a single set of routes. DiffServ, TOS marking, or other IP QoS mechanisms are used to select which of the two or more PVCs to use in either direction. Currently PVC bundles apply only to routed, not bridged interfaces. For them to be useful to this architecture, the approach would need to support bridged interfaces in addition to routed interfaces, and would need support simultaneous transport of both PPPoE and IP over one of the PVCs in the bundle.
Regional Network	The Regional Network interconnects between the Network Service Provider's network and the Access Network. A Regional Network for DSL connects to the BRAS, which is technically both in the Regional Network and in an Access Network. Typically more than once Access Network is connected to a common Regional Network. The function of the Regional Network in this document goes beyond traditional transport, and may include aggregation, routing, and switching.
Regional/Access Network	The Regional and Access Networks – grouped as and end-to-end QoS domain and often managed by a single provider.
Routing Gateway	A customer premises functional element that provides IP routing and QoS capabilities. It may be integrated into or be separate from the modem.

Session	A logically identifiable relationship formed between two (or more) communicating entities for exchanging control and data packets. An example of which would be a PPP session.
Subscriber	The client that is purchasing the DSL circuit from the Service Provider and is receiving the billing.
Traffic Classification	The process of selecting packets based on common criteria, such as the content of packet headers or session identification.
Traffic Marking	The process of setting packet header fields, such as DSCP, MPLS EXP or 802.1p/q COS field in a packet/frame/cell based on defined rules. Traffic marking may result from for example, a classification decision, a policing action, or a policy decision.
Traffic Metering	The process of measuring the rate and/or burst of a traffic stream selected by a classifier. The instantaneous state of this process may be used to affect the operation of a marker, shaper, or policer, and/or may be used for accounting and measurement purposes.
Traffic Policing	The process of dropping, marking or remarking packets/cells within a traffic stream in accordance with the state of a corresponding meter against a defined traffic profile, using mechanisms such as the token bucket scheme defined by [RFC2697].
Traffic Remarking	The process of changing header fields, such as DSCP, MPLS EXP or 802.1p/q COS field in a packet/frame based on defined rules.
Traffic Shaping	The process of delaying packets/cells within a traffic stream to cause it to conform to some defined traffic profile.
Traffic Stream	a set of one or more microflows or sessions, which are selected by a particular classifier.
Upstream	The direction of transmission from the modem to the Access Node.
User	Typically, a member, employee or guest at the Subscriber's household or business using the DSL circuit capabilities.

3.2 Broadband Provider Reference Definitions

Generally, services over a DSL access-based broadband network will be provided and supported by a number of different operational organizations. These organizations may be part of one company or more than one company and it is desirable to have a clear idea of the roles of the different organizations and how the functionality of equipment, network management, and test equipment can support their ability to discharge their roles for the benefit of the end customers. In order to provide a baseline with which to contrast, this document provides a common architectural view of DSL architecture in Figure 1.

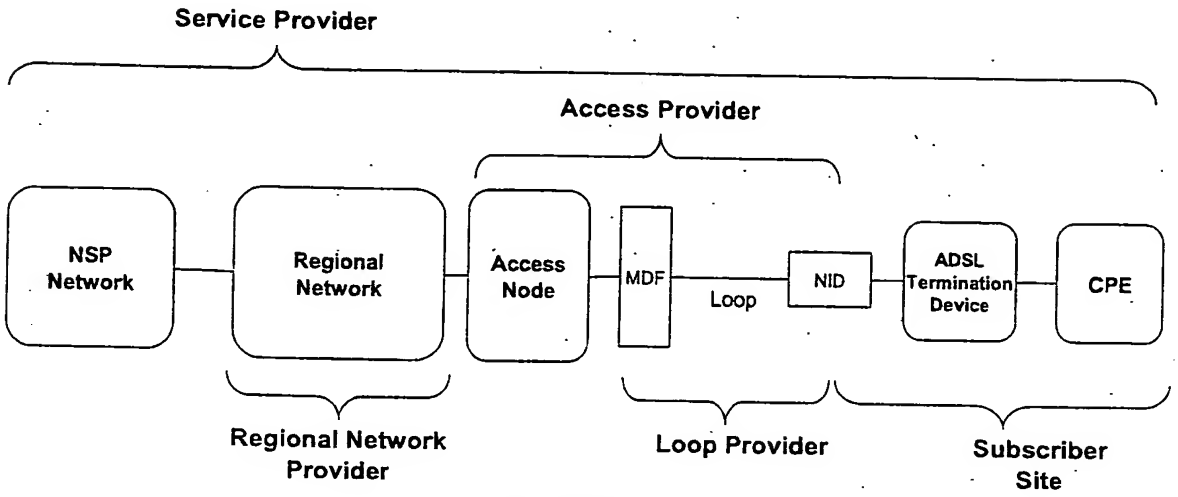

Figure 1 – DSL Network Components

(Voice components not shown for clarity)

Boxes in the figures represent functional entities – networks and logical components rather than physical elements.

This traditional architecture is centered on providing service to a line or a loop. It is desired, however, to be able to provide services that are user-specific. Additionally, more than one subscriber can be present at the same premises and share a single loop. There is a need, therefore, to describe a slightly more complex situation, and hiding the common complexity shared with Figure 1, this description is provided in Figure 2 below. Note that the figure shows many-to-many access through a common Regional/Access network. It is used to simultaneously provide an Application Service$_1$ between an ASP Network$_1$ and User$_1$ at the same time and over the same U interface as it supports a Network Service$_2$ between NSP Network$_2$ and User$_2$.

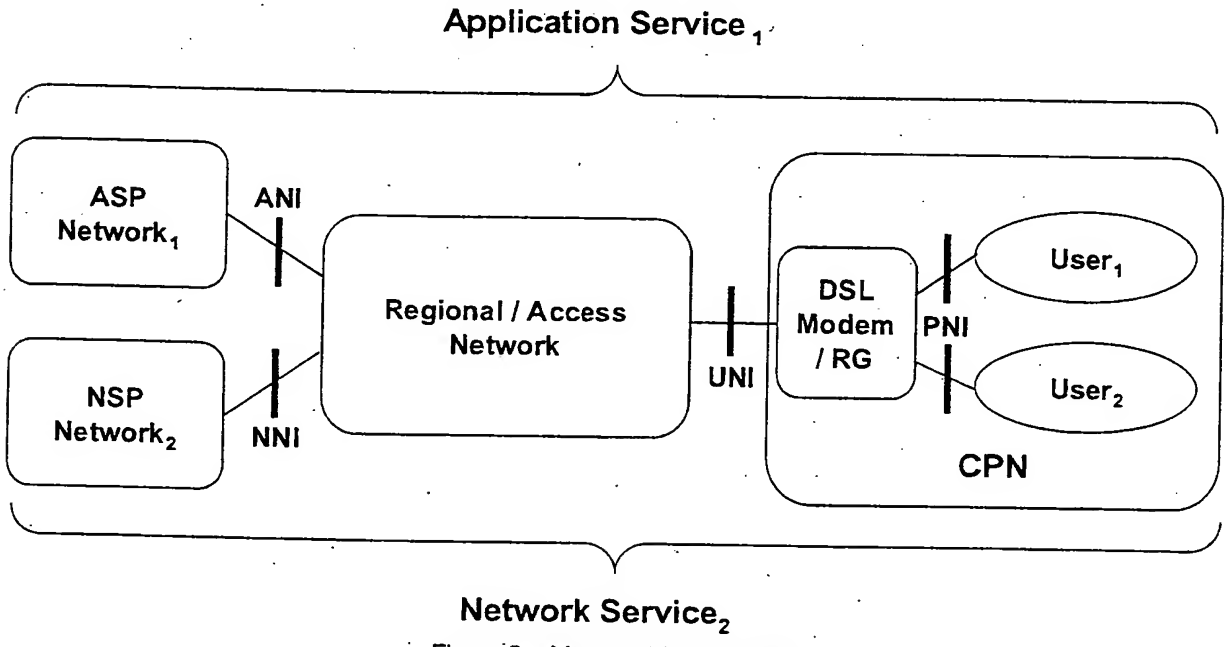

Figure 2 – Many-to-Many Access

The figures show the key components of a DSL access-based broadband network. They indicate ownership of the components to different providing organizations. The role of these various providers is indicated below:

The Network Service Provider (NSP):

- Includes Internet Service Providers (ISPs) and Corporate Service Providers (CSPs)
- Is responsible for overall service assurance
- May provide CPE, or software to run on customer-owned CPE, to support a given service
- Provides the customer contact point for any and all customer related problems related to the provision of this service
- Authenticates access and provides and manages the IP address to the subscribers

The Application Service Provider (ASP):

- Provides application services to the application subscriber (gaming, video, content on demand, IP Telephony, etc.)
- Is responsible for the service assurance relating to this application service
- Responsible for providing to subscribers additional software or CPE which specific services may require.
- Provides the subscriber contact point for all subscriber problems related to the provision of specific service applications and any related subscriber software.
- Does not provide or manage the IP address to the subscribers

The Loop Provider:

- Provides a metallic loop from the Access Network equipment to the customer's premises
- Is responsible for the integrity of the metallic loop and its repair
- May also provide the Access Network Provider aggregated access to remotely deployed DSL equipment owned, operated, and maintained by the Loop Provider

The Access Network Provider:

- Provides digital connectivity to the customer via the metallic Loop
- Is responsible for the performance and repair of the access transmission equipment

The Regional Network Provider:

- Provides appropriate connectivity between the Access Network and the NSPs and ASPs
- Is responsible for Regional Network performance and repair
- May perform aggregation services to NSPs or ASPs and/or may provide any to any connectivity within the RBN on behalf of the NSP/ASP.

3.3 Interfaces

These interfaces are key to this architecture, and have been modified or expanded from historical architectures (except the U interface) and represent requirements specific to the service models detailed herein and in WT-080.

3.3.1 A10-ASP Interface

This reference point is between the Regional/Access Network and the ASP's Points of Presence (POPs). This interface will consist of a routed IP interface, that may be transported over Fast Ethernet, Gigabit Ethernet, Packet over SONET (POS), or some other IP interface. The ASP has the end-to-end Service responsibility to the customer for their specific application and is viewed as the "Retailer" of the specific service. Trouble reports for the specific service go directly to the ASP.

3.3.2 A10-NSP Interface

This reference point is between the Regional/Access Network and the NSP's POPs. The interfaces could be ATM, Fast Ethernet, Gigabit Ethernet, or Packet over SONET (POS). In the case of ATM, multiple PPP sessions may be multiplexed over a single VCC at this interface. Typically, the NSP has the end-to-end service responsibility to the customer and is viewed as the "Retailer" of the service. As the retailer of the DSL service, trouble reports, and other issues from the subscriber are typically addressed to the NSP. Handoff protocols could include layer 2 (e.g. ATM VP/VCs, L2TP tunnels) and layer 3 (e.g. IPv4, IPv6 routed protocols).

3.3.3 U Interface

The U Interface is located at the subscriber premise between the Access Node and the DSL modem.

3.3.4 T Interface

The T Interface defines the interworking between the DSL modem/Routing Gateway and other CPE in the Customer Premises Network (CPN). The requirements for new vertical services over DSL require the addition of a Routing Gateway as the intermediate point between the DSL modem and the LAN Devices. The primary goal of this interface is to facilitate seamless transmission of IP packets in both a best effort approach as well as maintaining predefined QoS behavior or establishing dynamic QoS behaviors through a signaling mechanism. The DSL modem and Routing Gateway functions may or may not be combined in a single device.

4 REFERENCE ARCHITECTURE

4.1 Logical Reference Architecture

As noted in Section 3.2 above, the end-to-end DSL network consists of four providers. Of these providers, the two that this proposal most affects are the Regional Network Provider and the Access Network Provider. Historically the Regional Network has been a network of ATM switches, as shown in Figure 3. This is because the access to most Access Nodes is an ATM based interface. Some Access Networks even have their own ATM switches used to aggregate traffic from multiple Access Nodes.

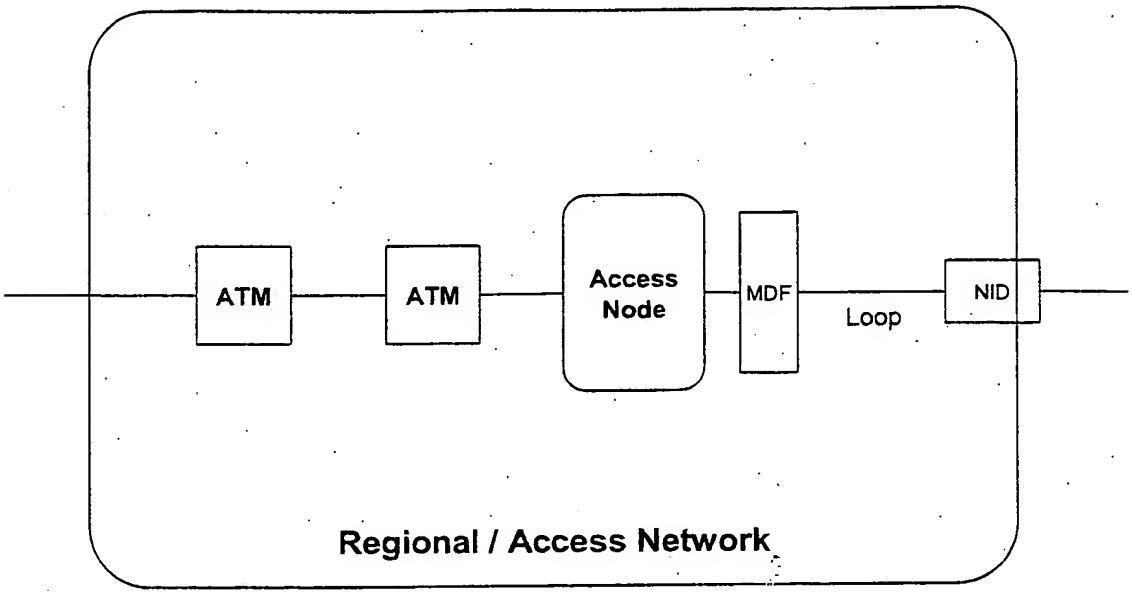

Figure 3 - ATM based Regional and Access Network Providers

In this architecture, there are no mechanisms for limiting subscriber traffic except for per line profiles within the Access Nodes. As many DSL networks were deployed before the advent of the BRAS, almost all the Access Network Providers use fixed speed profiles in the Access Nodes to limit upstream and downstream traffic. Even if the Service Provider were to attempt to send more traffic into the Regional Network than the Access Node is set to permit, the Access Node will police the downstream traffic. Since most Internet-based applications use TCP as the transport protocol, the traffic rejected at the Access Node will trigger TCP back-off, effectively throttling the downstream bandwidth. As such, most Service Providers also shape downstream traffic at the subscriber-selected bandwidth. However, the desire to move to a rate adaptive bandwidth model means that both the Regional and Access Networks could be vulnerable to traffic overloading. A means to control upstream and downstream traffic is needed as this architecture evolves.

Many times the physical components of the Access Nodes are daisy-chained, sharing the bandwidth of the aggregating circuit. As shown in Figure 13 in Section 4.2.5.4, there are numerous ways that DSL access devices can be interconnected to the first ATM switch. While historical measurements have shown that the typical DSL subscriber uses no more than a small fraction of sustained bandwidth, the fact is that as subscribers are offered more and more high bandwidth applications, the average sustained bandwidth per subscriber over these "mid-mile" connections is going to increase. As per subscriber bandwidth usage increases, the Regional Network Provider will also need to scale bandwidth and provide subscriber-level granularity. ATM VPs do not provide the granularity necessary to offer per application QoS on a per subscriber basis.

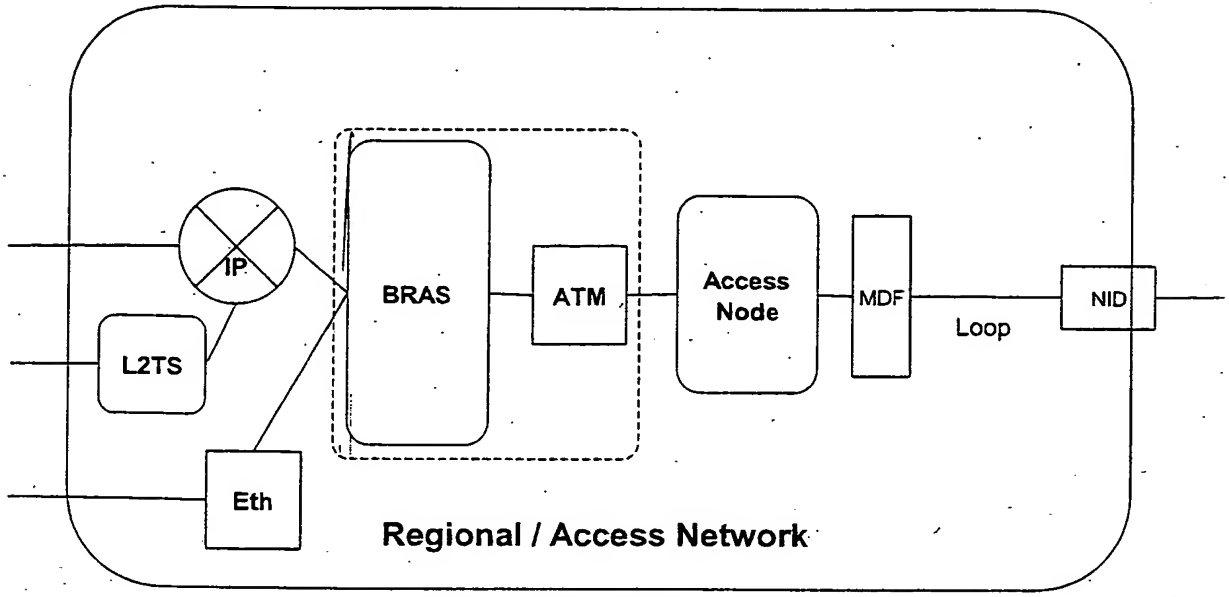

Figure 4 - IP Enabled Regional Network

As a result, other devices need to be added to the Regional Network to provide better aggregation of subscriber traffic. There are several options for doing this, most of which involve IP enabling the Regional Network as shown in Figure 4. Subscribers that use native IP, which is a routable protocol, can be aggregated at the IP level into a Virtual LAN (VLAN) or Virtual Private Network (VPN) for handoff to their associated Service Provider. Those subscribers that use variations of the Point-to-Point Protocol (PPP), such as PPPoA (PPP over ATM) and PPPoE (PPP over Ethernet), can be aggregated at either the PPP or the IP layer.

If the aggregation is done at the PPP layer, then these PPP sessions will need to be forwarded over a routable protocol such as Layer 2 Tunneling Protocol (L2TP). When the new subscriber aggregation element is functioning in this mode, it is referred to as an L2TP Access Concentrator or LAC. The other option for PPP based subscriber is to also terminate the PPP session and assign IP addresses to the subscribers. This traffic would then be collected into a VLAN or VPN as with native IP traffic. When performing PPP Termination and Aggregation (PTA), the box is typically called a Broadband Remote Access Server or BRAS.

As more and more DSL aggregation is performed at the IP layer rather than the ATM layer, additional transport options may be added. In addition to ATM, Ethernet and Packet over SONET are also options for IP transport. There are various metropolitan Ethernet solutions available in speeds of 10 Mbps (Ethernet), 100 Mbps (Fast Ethernet), or 1 Gbps (Gigabit Ethernet or GigE).

These new network elements also need to be able to function as the first tier ATM aggregation device, where the Access Node is now directly connected. As such, these devices will also need to handle ATM level aggregation and switching and need to function as an adjunct to the existing ATM network. Since they are IP aware, they can also serve as the Label Edge Router (LER) that is required if the Core Network is to become Multi Protocol Label Switching (MPLS) aware. This would be shown in Figure 4 by collapsing the BRAS and ATM switch into a single multi-protocol device.

4.2 Logical Elements and Interfaces

4.2.1 Application Service Provider Network

4.2.1.1 Description

The Application Service Provider (ASP) is defined as a Service Provider that uses a common infrastructure provided by the Regional/Access Network and an IP address assigned and managed by the Regional Network Provider. This is a new type of DSL service. The Regional Network Provider owns and procures addresses that they, in turn, allocate to the subscribers. ASPs then use this common infrastructure in order to provide

application or network services to those subscribers. For example, an ASP may offer gaming, Video on Demand, or even filtered Internet access, or access to VPNs via IPsec or some other IP-tunneling method. The ASP service may be subscriber-specific, or communal when an address is shared using Network Address Port Translation (NAPT) throughout a Customer Premises Network (CPN). It is envisioned that the ASP environment will have user-level rather than network-access-level identification, and that a common Lightweight Directory Access Protocol (LDAP) directory will assist in providing user identification and preferences. Logical elements used by ASPs typically include routers, application servers, and directory servers. The relationship between the ASP Network, the A10-ASP interface, and the Regional Network is shown in Figure 2. There is one and only one ASP network per Regional/Access Network.

4.2.1.2 Capabilities

The capabilities of the ASP include but are not limited to the following:

- Authenticating users at the CPN
- Assignment of user profile or preference data
- Assignment of QoS to service traffic
- Customer service and troubleshooting of network access and application-specific problems
- Ability to determine traffic usage for accounting purposes and billing

4.2.2 A10-ASP Interface

4.2.2.1 Functionality

As shown in Figure 5, the A10-ASP interface defines the interworking between the ASP Network and the Regional/Access Network. This is not a traditional interface. However, in order to provide more technical and business options to would-be broadband content and application providers this document defines a way for a Service Provider to attach a server, servers, or entire network to a common infrastructure directly accessible by DSL subscribers. The A10-ASP interface is intended to promote content on demand, IP telephony, gaming, and other Quality of Service (QoS) or Bandwidth on Demand (BoD) applications without the need to deploy or manage an IP infrastructure. This also conserves IP addresses, as a single address can be used to gain access to all the services and providers that opt to share this infrastructure.

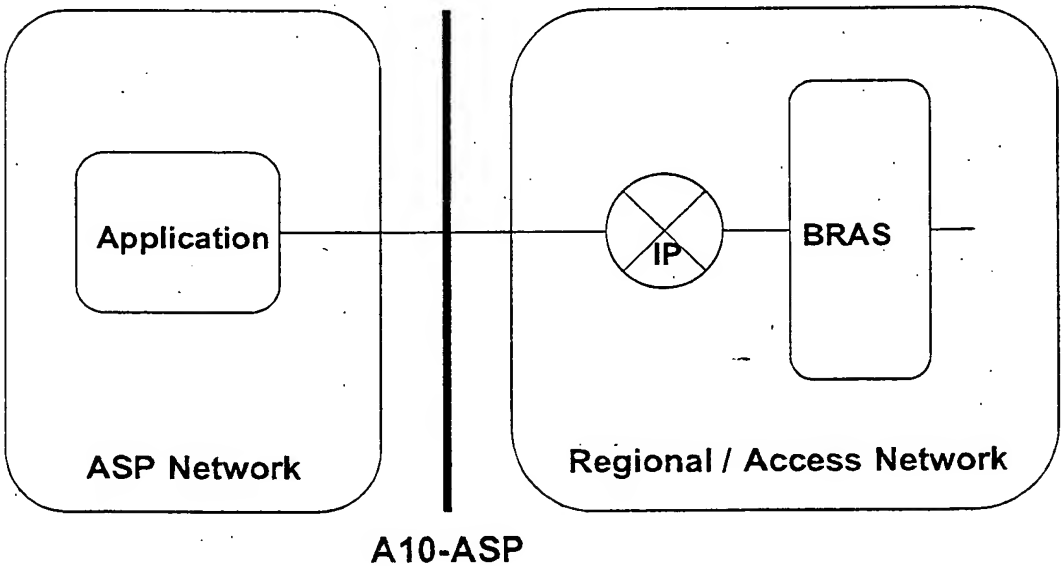

Figure 5 - A10-ASP Interface

4.2.2.2 Communication Protocols

This interface MUST[1] support IP networking connectivity to the DSL subscribers. Several QoS and BoD use cases exist:

1. Best effort IP networking is used with no additional QoS or information required.
2. Differentiated Services (Diffserv) QoS is provided in order to establish a higher class of service – oriented toward higher throughput, packet precedence, or lower latency.
3. QoS and Bandwidth limitations can be enforced by the Regional/AccessNetwork based on provisioned relationships between the ASP and all users or potentially specific users.

The communications protocol stack is shown in the following Figure 6.

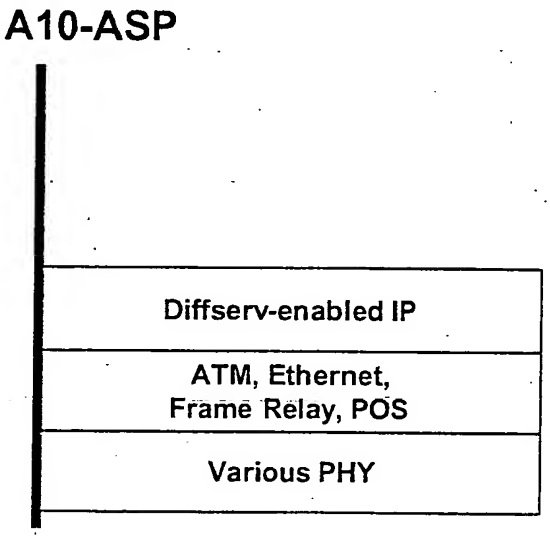

Figure 6 - ASP Protocol Stack with QoS

The ASP obtains an IP connection over a typical data link layer as described earlier. More likely is that an ASP actually obtains a 10 Base-T, 100 Base-T, or GigE connection to the Regional/Access Network within a co-location or hosting facility. The Regional/Access Network provider statically assigns addresses to the A10 ASP interfaces, and MAY[2] provide address blocks to the ASP.

Network Layer

The network layer interface MUST[3] support IP version 4 in accordance with IETF RFCs 791 and 2474.

The network layer MAY[4] support IP version 6 in accordance with IETF RFC 2460.

The network layer interface SHOULD[5] support IP multicast.

The network layer interface MUST[6] support IP precedence based on Diffserv Code Point (DSCP) markings, in accordance with IETF RFC 3140 when that type of QoS is offered. In other words, IP QoS will use Diffserv instead of using TOS bits or other potential indicators and definitions.

Data Link Layer

The data link layer SHOULD[7] support Ethernet in hosting or co-location sites.

The data link layer MAY[8] support ATM, Frame Relay, and/or POS.

The data link layer MAY[9] support bonding of multiple physical interfaces.

Physical Layer

The physical layer interface MUST[10] support at least one of the following – as appropriate:

- Ethernet PHY for 10 Mbps, 100 Mbps, 1 Gbps
- DS1, DS3, E1, E3
- OC3c, OC12c, OC48c, STM1c, STM4c, STM16c

4.2.3 Network Service Provider Network

4.2.3.1 Description

The Network Service Provider (NSP) is defined as a Service Provider that provides addressing and connectivity to an Internet Protocol (IP) network. This is the typical application of DSL service today. The NSP owns and procures addresses that they, in turn, allocate individually or in blocks to their subscribers. The subscribers are typically located in Customer Premises Networks (CPNs). The NSP service may be subscriber-specific, or communal when an address is shared using NAPT throughout a CPN. This relationship among the NSP, A10-NSP interface, and Regional/Access Network is shown in Figure 2. NSPs typically provide access to the Internet, but may provide access to a walled garden, VPN, or some other closed group or controlled access services. L2TP and IP VPNs are typical A10-NSP interface arrangements.

The capabilities of the NSP include but are not limited to the following:

- Authenticating network access between the CPN and the NSP network
- Assignment of network addresses and IP filters
- Assignment of traffic engineering parameters
- Customer service and troubleshooting of network access problems

4.2.4 A10-NSP interface

4.2.4.1 Functionality

As shown in Figure 7 and Figure 9, the A10-NSP interface defines the interworking between the NSP and the Regional/Access Network provider. This document offers the following Layer 2 and Layer 3 options for this interconnection.

4.2.4.2 Communication Protocols: L2TP Connection

This interface $MUST_{[11]}$ support the Layer 2 PPP connection service supported by L2TP. Using Figure 8 as a reference, subscribers $MUST_{[12]}$ be placed into L2TP tunnels in one of the following methods:

1. L2TP tunnels $MAY_{[13]}$ be established or provisioned statically between LNS and the LAC or through an intervening Layer 2 Tunnel Switch (L2TS).
2. L2TP tunnels $MAY_{[14]}$ be established dynamically using RADIUS in order to determine which users to add to various L2TP tunnels, including potentially new ones. As before, these may be directly between LAC and LNS or via L2TS.

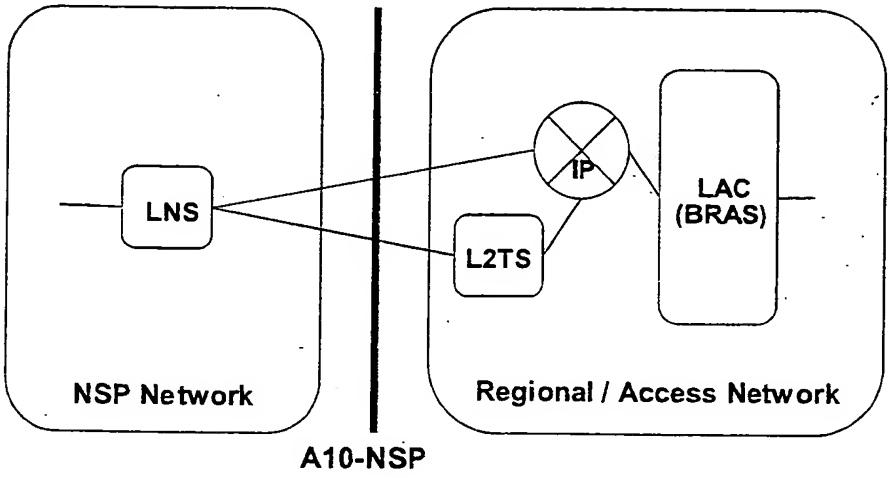

Figure 7 - A10-NSP Interface Supporting L2TP Connection

One or more concurrent sessions can be established to NSPs from any given CPN, and the destinations are chosen by the fully qualified domain name (FQDN) of the accessing subscriber.

Business models that require limiting subscriber access to a single NSP SHOULD[15] be supported through administrative limits on the FQDN routing established by the Regional/Access Network provider on behalf of one or more NSPs. Subscribers SHOULD[16] be able to establish multiple access sessions to the same or to different NSPs.

The RADIUS response MAY[17] be used to determine the bandwidth profile for its access session. Note that RADIUS will require enhancement to do this in a standard way.

The communications protocol stack is shown in the Figure 8.

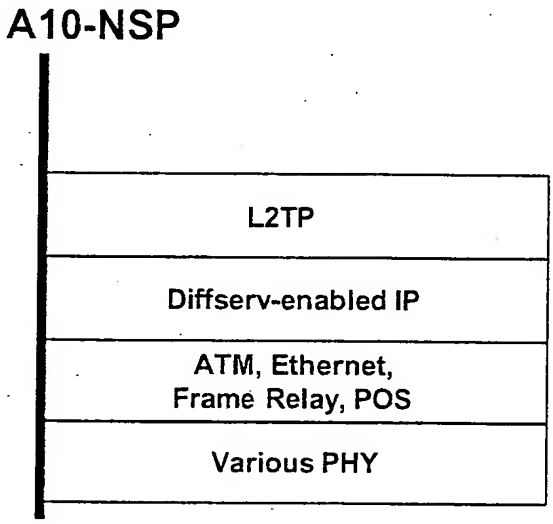

Figure 8 - L2TP Protocol Stack

While L2TP over IP is always used, as opposed to L2TP delivered directly over ATM or Frame Relay, various IP transport options can be provided by the Regional/Access Network provider or selected by the NSP according to availability, regulation, and economics. Also, while the entire L2TP tunnel can be provided with a traffic engineering specification, the constituent flows within an L2TP tunnel will not receive differentiated service. In other words all the flows within an L2TP tunnel will receive the same aggregate QoS treatment.

Network Layer

The network layer interface MUST[18] support IP version 4 in accordance with IETF RFCs 791 and 2474.

The network layer MAY[19] support IP version 6 in accordance with IETF RFC 2460.

The network layer MUST[20] make use of L2TP over IP in accordance with IETF RFC 2661.

Data Link Layer

The data link layer SHOULD[21] support ATM.

The data link layer MAY[22] support Ethernet, Frame Relay, and/or POS.

The data link layer MAY[23] support bonding of multiple physical interfaces.

Physical Layer

The physical layer interface MUST[24] support at least one of the following – as appropriate:

- Ethernet PHY for 10 Mbps, 100 Mbps, 1 Gbps
- DS1, DS3, E1, E3
- OC3c, OC12c, OC48c, STM1c, STM4c, STM16c

4.2.4.3 Communication Protocols: IP Routed Connection

This interface MUST[25] support the Layer 3 IP routed connection. Using Figure 9 as a reference, subscribers MUST[26] be placed into IP routed networks in one of the following methods:

1. IP address pools MAY[27] be established or provisioned statically.
2. IP addresses MAY[28] be provided in pools that are distributed dynamically by the Regional/Access Network provider.
3. Subscribers IP addresses MAY[29] be distributed from the NSP to the BRAS dynamically using RADIUS.
4. IP addresses MAY[30] be assigned from named pools in cases where the NSP opts to allocate addresses out of two or more pools based on subscriber-specific information.

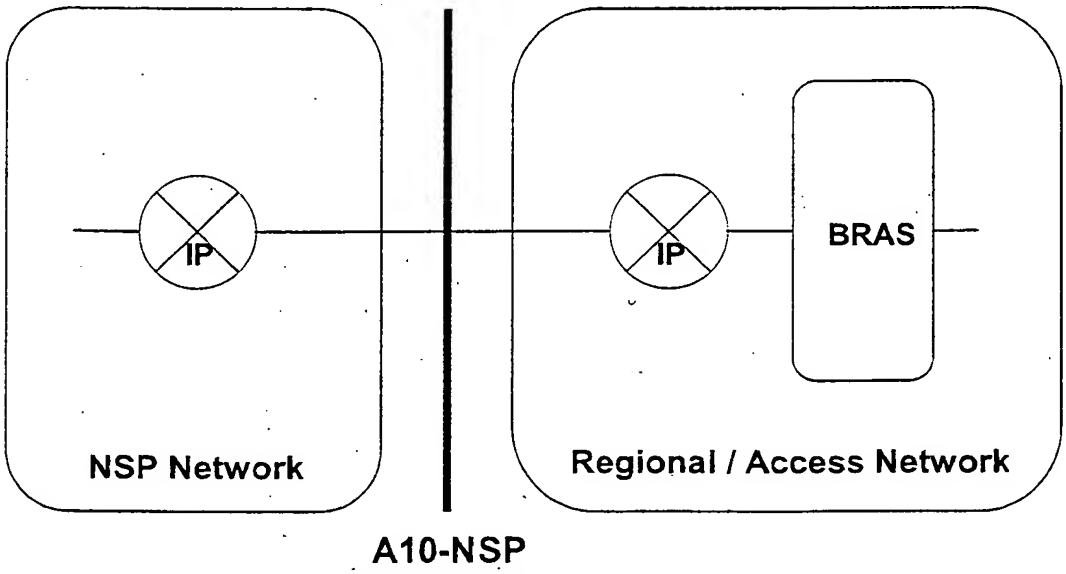

Figure 9 - A10-NSP Interface Supporting IP Routed Connection

In every case, RADIUS MUST[31] be used between the BRAS (or a potential RADIUS proxy) and an NSP-designated AAA system or systems to authenticate subscriber access to the routed network.

In most cases, the IP routed network will be comprised of many IP-VPNs that support sharing of the Regional/Access Network at the IP layer.

Multiple services may be offered across the 'U' interface. Access to a particular IP service will be established as with L2TP using the Network Access Identifier (NAI a.k.a. FQDN) provided by the accessing subscriber. Subscribers MUST[32] be able to establish multiple access sessions to the same or to different NSPs. Business models that require restricting simultaneous access to particular combinations of IP service MUST[33] be supported through administrative policies established in the regional/access network on behalf of the NSPs/ASPs.

If an NSP connects to the Regional/Access Network in several places, the A10-NSP interface SHOULD[34] support BGP4 as per IETF RFC 1745.

Several QoS and BoD use cases exist:

1. Best effort IP networking is used with no additional QoS or information required.

2. Diffserv QoS MAY[35] be supported and MAY[36] be used in order to establish a higher class of service – oriented either toward higher throughput, or lower latency.

3. The Regional/AccessNetwork can enforce QoS and Bandwidth limitations based on provisioned relationships between the NSP and all users or potentially specific users.

The communications protocol stack is shown in Figure 10.

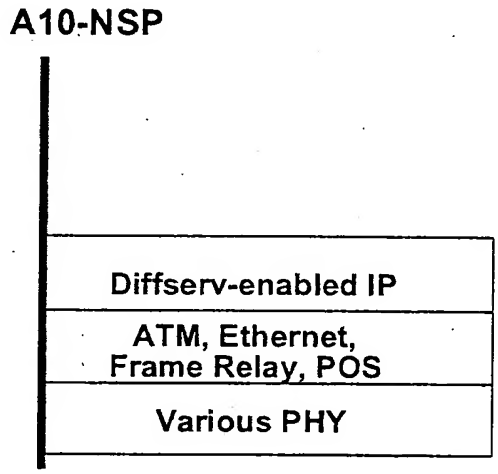

Figure 10 – Routed IP Protocol Stack with QoS

IP MUST[37] always be used; however, various IP transport options can be provided by the Regional/Access Network provider or selected by the NSP according to availability, regulation and economics. As described earlier, RADIUS MUST[38] always be used to authenticate users, SHOULD[39] be used to set NSP-desired filters, and MAY[40] be used to assign addresses.

Network Layer

The network layer interface MUST[41] support IP version 4 in accordance with IETF RFCs 791 and 2474.

The network layer interface MUST[42] support IP precedence based on Diffserv Code Point (DSCP) markings, in accordance with IETF RFC 3140 when this type of QoS is offered.

The network layer interface SHOULD[43] support IP multicast.

The network layer interface MUST[44] support IP precedence based on Diffserv Code Point (DSCP) markings, in accordance with IETF RFC 3140 when this type of QoS is offered.

The network layer MAY[45] support IP version 6 in accordance with IETF RFC 2460.

Data Link Layer

The data link layer SHOULD[46] support ATM

The data link layer MAY[47] support Ethernet, Frame Relay, and/or POS.

The data link layer MAY[48] support bonding of multiple physical interfaces.

Physical Layer

The physical layer interface MUST[49] support at least one of the following – as appropriate:

- Ethernet PHY for 10 Mbps, 100 Mbps, 1 Gbps
- DS1, DS3, E1, E3
- OC3c, OC12c, OC48c, STM1c, STM4c, STM16c

4.2.5 Regional/Access Network

The Regional/Access Network consists of the Regional Network, Broadband Remote Access Server, and the Access Network as shown in Figure 11. Its primary function is to provide end-to-end transport between the customer premises and the NSP or ASP. The Regional/Access Network may also provide higher layer functions such as QoS and content distribution. QoS will be provided by tightly coupling traffic-engineering capabilities of the Regional Network with the capabilities of the BRAS. Depending on the type and frequency of use, certain content storage may be pushed further out in the Regional/Access Network than others. As a result, functionality to support content distribution could be located at different points within the Regional/Access Network, but will not be located between the BRAS and the subscriber.

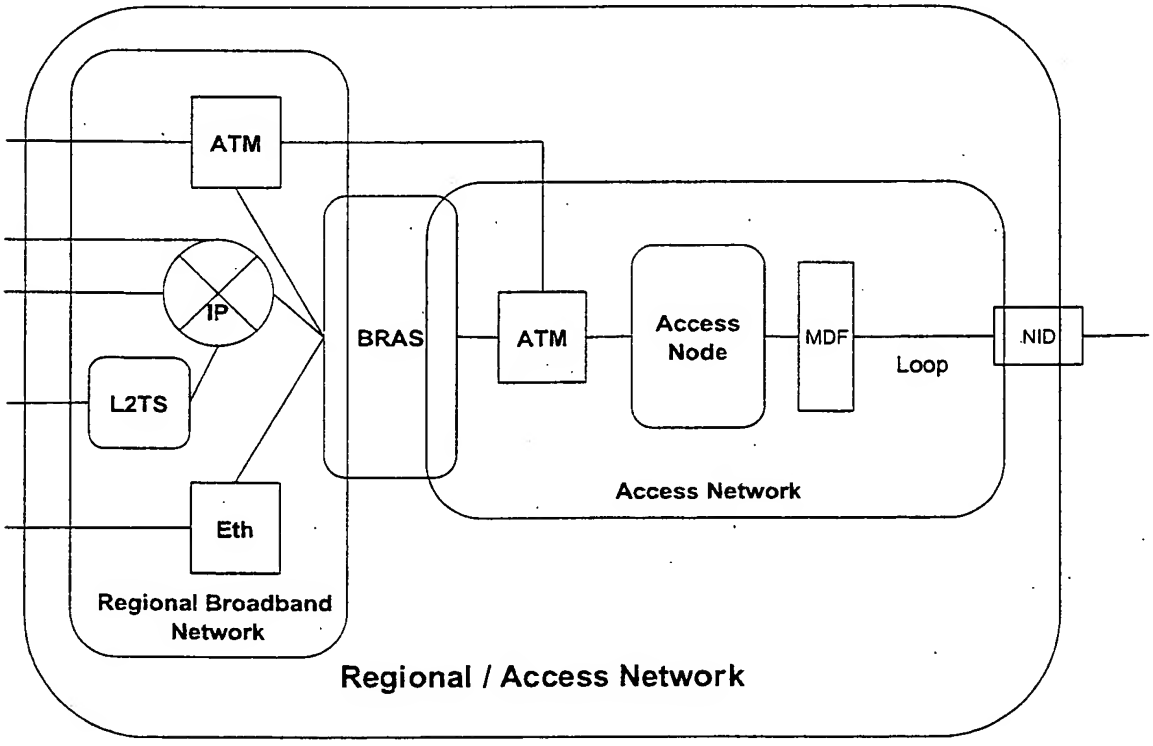

Figure 11 - Components of the Regional/Access Network

4.2.5.1 Regional Network

The Regional Network connects one or more BRAS and associated Access Network to NSPs and ASPs. It supports aggregation of traffic from multiple Access Networks and hands off larger geographic locations to NSPs and ASPs – relieving a potential requirement for them to build infrastructure to attach more directly to the various Access Networks. This arrangement is shown in Figure 12, which pictures an NSP and an ASP attached to a Regional Network in order to gain access to 3 Access networks. This architecture assumes that the network providers of the Regional and Access Networks work extremely closely in order to provide an end-to-end QoS solution. A good assumption might be that the 2 networks are operated and managed by a single service providing entity and offered as a combined, Regional/Access Network.

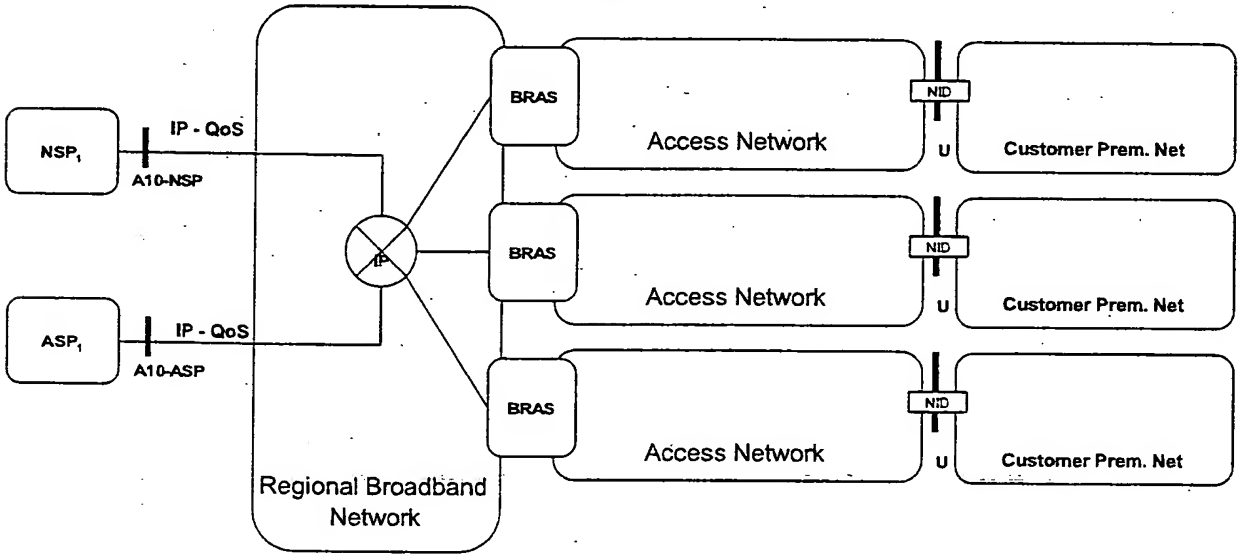

Figure 12 - Aggregation function of Regional Network

The Regional Network may transport traffic using ATM, Ethernet, IP or MPLS. Within these networking technologies, the Regional Network MUST[50] provide scalable traffic engineering capabilities and preserve IP QoS.

4.2.5.2 Broadband Remote Access Server

The BRAS performs multiple functions in the network. Its most basic function is to provide aggregation capabilities between the Regional/Access Network and the NSP/ASP. For the aggregation Internet traffic, the BRAS serves as a L2TP Access Concentrator (LAC) tunneling multiple subscriber PPP sessions directly to an NSP or switched through a L2TS. It also performs aggregation for terminated PPP sessions or routed IP session by placing them into IP VPNs or 802.1Q VLANs. The BRAS also supports ATM termination and aggregation functions.

Beyond aggregation, the BRAS is also the injection point for providing policy management and IP QoS in the Regional and Access Networks. The BRAS is fundamental to supporting the concept of many-to-many access sessions.

Policy information can be applied to terminated and non-terminated sessions. For example, a bandwidth policy may be applied to a subscriber whose PPP session is aggregated into an L2TP tunnel and is not terminated by the BRAS. However, sessions that terminate on (or are routed through) the BRAS can receive per flow treatment because the BRAS has IP level awareness of the session. In this model, not only can the aggregate bandwidth for a customer be controlled but also the bandwidth and treatment of traffic on a per application basis.

The delivery of content has shifted from content that was more download intensive with lower bandwidth and best effort quality to one that is more real-time in nature, requiring higher bandwidth with higher quality. Some of the higher bandwidth applications include Video on Demand (VoD) for movies, multicast ("Broadcast" TV), and MPEG unicast video. Given the BRAS's proximity to the edge of the network and its ability to support IP

services, the BRAS SHOULD[51] also provide support for content distribution and efficient use of multicast services.

Some high level functional requirements are for the BRAS are listed below. This list is not comprehensive and additional requirements for QoS are listed in Section 5. Additionally, BRAS requirements from both this architecture as well as other architectures are expected to become a separate DSL Forum topic.

- The BRAS MUST[52] be able to aggregate PPP sessions into L2TP tunnels (LAC function).
- The BRAS MUST[53] be able to terminate PPP sessions and assign routing attributes based on subscriber profile (LNS function).
- The BRAS MUST[54] support authentication using RADIUS.
- The BRAS MUST[55] support IP over bridged Ethernet (IETF RFC 2684).
- The BRAS MUST[56] support address allocation using Dynamic Host Configuration Protocol (DHCP).
- The BRAS MUST[57] support multiple VCs per subscriber.
- The BRAS SHOULD[58] support ATM VC/VP cross-connection functions independent of AAL type.
- The BRAS MUST[59] support termination and aggregation of ATM VCs.
- The BRAS SHOULD[60] support the following ATM classes of service: UBR, UBR+, CBR, VBR-nrt, VBR-rt .
- The BRAS MUST[61] allocate downstream bandwidth based on policy configuration across ATM, PPP, Ethernet, and IP technologies.
- The BRAS MUST[62] mark IP QoS fields for upstream and downstream traffic based on policy configuration.
- The BRAS MUST[63] support policing of upstream per-subscriber traffic based on policy configuration.
- The BRAS MUST[64] support queuing and prioritization based on diffserv marking and/or flow classification.
- The BRAS MUST[65] support traffic engineering for networking technologies including ATM, MPLS, and Ethernet.
- The BRAS MUST[66] support a Diffserv-aware hierarchical scheduler that allows it to manage the network so that any potential congestion in the Access Network between the BRAS and the RGs is avoided. The hierarchical scheduler in the BRAS MUST[67] be able to model the congestion points in at least two subsequent ATM hops (corresponding to the daisy chaining of two ATM switching/multiplexing points in the Access Node); if the BRAS does not include the ATM switching function, then the hierarchical scheduler in the BRAS MUST[68] be able to model the congestion point in yet an additional ATM hop. This scheduler is described in further detail in section 5 and shown by example in Appendix B.
- The BRAS MUST[69] shape the individual subscriber's aggregate downstream traffic to the subscribed rate which will be some value equal to or lower than the DSL sync rate.
- The BRAS MUST[70] support RED and WRED policing of upstream traffic using the same topology information that exists for the hierarchical scheduler.
- When operating in an IP-routed mode, the BRAS MAY[71] provide multicast support
- The BRAS SHOULD[72] support Ethernet LAN interfaces for the local attachment of content distribution servers.
- When operating in an IP-routed mode the BRAS MAY[73] provide multicast access control and collect multicast usage information.

4.2.5.3 Access Network

Description

The Access Network refers to the network between the NID and the BRAS. The protocols between these devices are well defined and this recommendation does not attempt to alter them.

4.2.5.4 Access Node

Description

The Access Node contains the XTU-C, which terminates the DSL signal. Physically, the XTU-C can be deployed in the central office in a DSLAM, or remotely in a remote DSLAM (RT-DSLAM), Next Generation Digital Loop Carrier (NG-DLC), or a Remote Access Multiplexer (RAM). A DSLAM hub can be used in a central office to aggregate traffic from multiple remote physical devices, and is considered logically to be a part of the Access Node.

The Access Node provides aggregation capabilities between the Access Network and the Regional Network. It is the first point in the network where traffic on multiple DSL lines will be aggregated onto a single network. Traditionally the Access Node has been primarily an ATM concentrator, mapping PVCs from the DSL modem to PVCs in the ATM core. It has also shaped and policed traffic to the service access rates.

The role of the Access Node will change slightly in this architecture. While it will remain in the aggregation role, the current responsibility of policing DSL modem-to-BRAS PVCs to the subscribed line rate will be moved from the Access Node to the BRAS in order to establish additional bandwidth on the DSL line for additional services. The Access Node will set line rate for each PVC at the synch rate (or slightly less) of the DSL Modems. This will make the maximum amount of subscriber bandwidth available for services. The BRAS will retain the ability to police individual sessions/flows as required to their existing rates and will also perform the dynamic changes when bandwidth-on-demand services are applied. In order to do this the BRAS MUST$_{[74]}$ be provisioned so that it does not allow traffic to flow faster than the DSL sync rate. The BRAS MAY$_{[75]}$ be provisioned with the actual DSL sync rate to accomplish this.

Various physical Access Node configurations are shown in Figure 13, with brief names for the configurations listed in Table 1.

In order to allow IP QoS support over an existing non-IP-aware layer 2 network without using multiple layer 2 QoS classes, a mechanism based on hierarchical scheduling is used. This mechanism, which is further described in section 5, preserves IP QoS over the ATM network between the BRAS and the RGs by carefully controlling downstream traffic in the BRAS, so that significant queuing and congestion does not occur further down the ATM network. This is achieved by using a hierarchy of scheduling steps in the BRAS that will account for downstream trunk bandwidths and DSL synch rates. As the depth of non-IP aware nodes between the BRAS and RG increases, the complexity of implementing hierarchical scheduling grows as well. In order to minimize this complexity, the daisy chaining MUST NOT$_{[76]}$ exceed a depth of more than two ATM switching / multiplexing points including the Access Node and subtending Access Nodes. Additionally, if the BRAS does not incorporate an ATM pass-through or switching functionality, an additional layer of hierarchical scheduling MUST$_{[77]}$ be used to manage the trunk to the ATM switch.

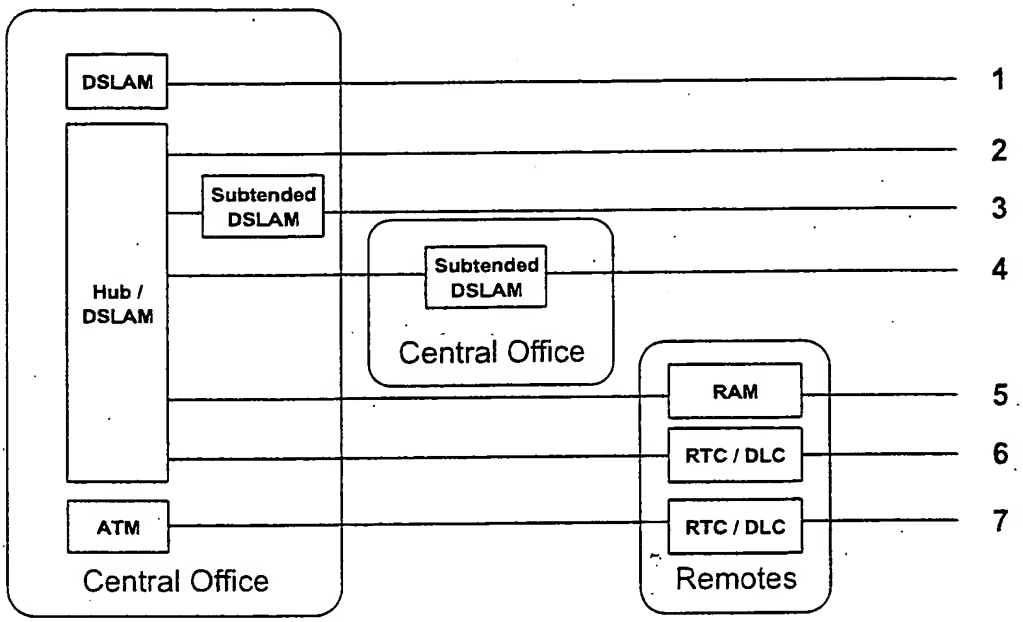

Figure 13 – Access Node Architecture Variations

Table 1 – Access Node Architecture Variation Descriptions

Reference #	Description
1	Access Node
2	Hub Access Node
3	Collocated Subtended Access Node
4	Remotely Located Subtended Access Node
5	Subtended Remote Access Node
6	Subtended DLC Located Access Node
7	Aggregated DLC Located Access Node

4.2.6 U Interface

4.2.6.1 Functionality

The U interface is defined as the interface between the Access Network and the CPN. This interface refers to the area between the CPN where the DSL modem is located and the Access Network where the Access Node is located – usually in the NID. The U interface includes the capabilities and protocols that cross between the Access Network and the CPN.

4.2.6.2 Communication Protocols

As shown in Figure 14 the U interface defines the interworking between the CPN and the Regional/Access Network. This interface MUST[78] support the bi-directional delivery of data for all the new product and service definitions as well as for existing (legacy) products and services.

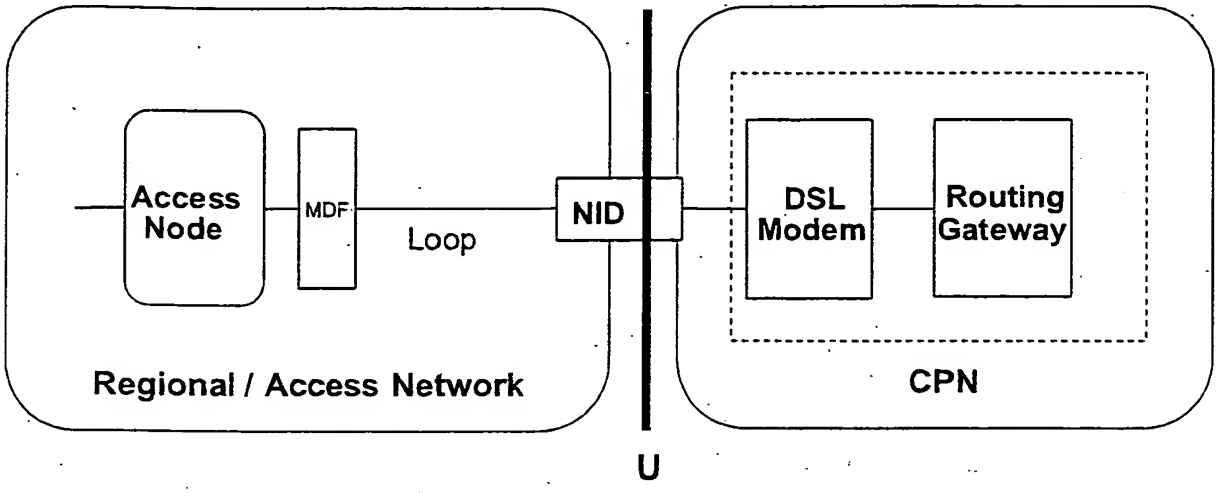

Figure 14 – U Interface

Although the first Network Element connection in the network is at the Access Node, the U interface MUST[79] support the transparent flow of protocols from the DSL Modem to the BRAS.

- The U interface MUST[80] support at least one ATM AAL5 PVC per CPN using PPPoE and/or IP over Ethernet (IETF RFC 2684 configured using DHCP). Although the target architecture to support QoS enabled IP services seeks to utilize a single ATM AAL5 PVC per CPN, it is recognized that certain required network element features identified in this document have yet to be developed. In particular, dynamic packet fragmentation/MTU sizing in the CPE (needed to control jitter and delay for short packet/high priority applications) may trail the availability of other required network element features. In order to meet the demands of service descriptions previously identified in an acceptable timeframe, a second ATM PVC may be provisioned to provide a means to separate those application flows having tight jitter and latency requirements. This second PVC will require that DSL modems support multiple PVCs. In the event that 2 PVCs are provisioned, it is desired that they be treated as a PVC bundle as this feature is made available. Additionally the PVC bundle standards need to be enhanced to support bridged Multi-service traffic.
- The U interface MUST[81] support Diffserv Code Points (DSCP) per IETF RFCs 2474 and 3260, enabling application-layer QoS access.
- The U interface MUST[82] support the ability to dynamically push IP routes back to the customer PC or Routing Gateway. Thus, RIPv2 will be used to provide routes to the RG. The RG is not expected to provide routes to the WAN.
- The BRAS SHOULD[83] support a mechanism to push routing information to the RG at the start of a PPP session.

The communications protocol stack is shown in the following figure.

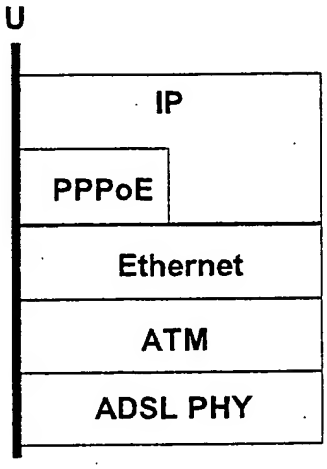

Figure 15 – U Interface Protocol Stack

Network Layer

The network layer interface MUST[84] support IP version 4 in accordance with IETF RFCs 791 and 2474.

The network layer MAY[85] support IP version 6 in accordance with IETF RFC 2460.

The network layer interface MAY[86] support IP precedence based on Diffserv Code Point (DSCP) markings, in accordance with IETF RFC 3140.

The network layer interface MUST[87] support PPPoE per IETF RFC 2516.

Data Link Layer

The data link layer MUST[88] support Ethernet encapsulation in accordance with IETF RFC 2684.

The data link layer MUST[89] support ATM in accordance with ATM Forum standards.

Physical Layer

The physical layer interface MUST[90] support G.dmt, and its related standards.

4.2.7 Customer Premises Network

The Customer Premises Network (CPN) is defined at its highest level as the location where the ATU-R is located and terminates the physical DSL signal, and where the subscriber's computers and other devices are interconnected. The initial DSL deployments focused on single user architectures where the CPN constituted a single PC connected directly to a DSL modem. This paradigm of service will continue to be supported and improved, but must be extended to support advanced features that go beyond the single user model. To support enhanced features (multi-user, gaming, VoIP, video, etc), the CPN must evolve to support the networking and management of devices and services within the home or business location.

From a network perspective, the CPN is the ultimate target of the services provided by the Service Provider (NSP or ASP). The CPN includes the networking environment and protocols that are resident in the premises. A CPN may imply coexistence of different link and physical layer technologies such as radio, power line transmission and Ethernet, but is assumed to have access to outside networks (via DSL). The terms devices and appliances refer to the collection of end terminals that can reside on the CPN, either temporarily (laptops, palm pilots, foreign devices etc.) or permanently, such as desktops, security, and climate control systems. Devices may or may not be individually addressable and reachable from other devices, inside or outside the CPN. Some devices may communicate with proxies that then can relay or translate state or configuration information for these end devices.

4.2.7.1 DSL Modem

Description

The DSL Modem terminates both DSL and ATM. It may or may not be integrated with additional Routing Gateway (RG) functionality. If it is not integrated, it will be used in a mode that is referred to as a simple bridge modem.

Capabilities

The capabilities of the DSL Modem in support of this architecture MUST[91] include but are not limited to the following:

- 2 ATM AAL5 PVCs - in order to be able to support the U interface service option of using 2 PVCs as described in 4.2.6.2. Note that in practice, DSL modems will likely have additional service drivers that would require them to support additional PVCs.
- UBR, UBR+ and VBR-rt ATM classes of service
- Per-VC queuing, separate priority queues for ATM classes of service

4.2.7.2 Routing Gateway

Description

CPN architectures typically leverage a Routing Gateway (RG) device that provides functionality beyond that of a basic DSL modem. The RG may or may not be integrated with the DSL modem function. In the integrated scenario, the device terminates the DSL signal from the network and provides an interface to other equipment located within customer premises. In the non-integrated case, the RG is physically separate from the DSL modem and adds functionality to the CPN independent of the DSL modem.

The principal tasks of the RG are to shape upstream traffic to the policed rate at the BRAS, to provide appropriate queuing and precedence for QoS traffic, and to allow a home network to share a single public address for network access. The data required for these duties may be pre-provisioned, user-provisioned or may be provisioned using an automatic configuration protocol. For an example of this third case, the RG may query a configuration server in the Regional / Access Network in order to learn the upstream policing rates for its access connections – and in the case of a non-integrated RG it may also learn the upstream sync rate of the detached ATU-R.

Since the integrated RG has knowledge of the CPN and its access to external networks, it enables tighter control of QoS for real time applications than may be possible in a non-integrated architecture. Both integrated and non-integrated RG are supported in this specification.

Capabilities

To support this QoS-enabled architecture, the capabilities of the RG MUST include but are not limited to the following:

- IP routing between the CPN and the Access Network [92]
- Multi-user, multi-destination support: Multiple simultaneous PPPoE sessions (started from the RG or from devices inside the CPN) in conjunction with non-PPP encapsulated IP (bridged) sessions per IETF RFC 2684. [93]
- Network Address Port Translation (NAPT) [94]
- Local DHCP [95]
- Support for major applications and protocols in the presence of NAPT and firewall (e.g., SIP, H.323, IPsec) [96]
- Dynamic MTU negotiation [97]
- Packet segmentation based on traffic/queue type [98]

- PPPoE pass though[99]
- Multiple queues, with the appropriate scheduling mechanism.[100]
- IP QoS
  - Classification, scheduling and shaping of IP flows[101]
  - Diffserv[102]
  - Management interface[103]
  - Support for real time services (Voice, Video)[104]
  - Re-marking capabilities[105]
- If 2 VCs are provisioned, support the mapping between Diffserv Code Point (DSCP) and a specific PVC (Using a PVC bundle is the desired way to meet this requirement)[106]

4.2.7.3 Networking Technologies

Description

The CPN will support the transparent transmission of IP packets. It is expected that the CPN will be a hybrid of technologies that may include Ethernet, phone line networking, power line networking, wireless networking, and others.

4.2.7.4 LAN Devices

Description

Devices inside the CPN that are served by the DSL Modem and RG, and connected by the various Networking Technologies are referred to as LAN Devices. These may include, but are not limited to, PCs, laptops, networked set-top boxes, and Internet Appliances.

4.2.8 T Interface

4.2.8.1 Functionality

As shown in Figure 16, the T interface defines the interworking between the DSL modem/RG and the LAN Devices. This interface MUST[107] support the bi-directional delivery of IP packets between the RG and other CPE as well as the ability to assign addresses to other CPE using DHCP. The other major functional requirement placed on the T interface includes identifying and supporting "QoS flows" as defined in Section 5. The primary goal of this interface is to facilitate seamless transmission of IP packets in both a best effort approach as well as maintaining predefined QoS behavior.

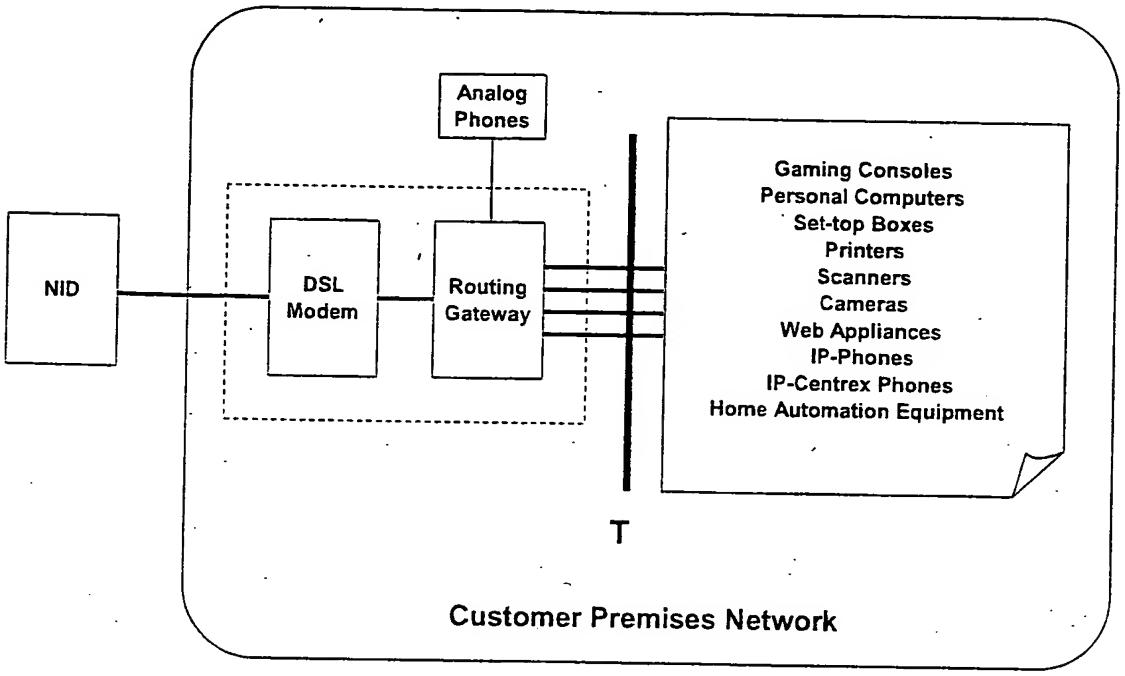

Figure 16 - T Interface

4.2.8.2 Communication Protocols

Network Layer

The network layer interface MUST[108] support IP version 4 in accordance with IETF RFCs 791 and 2474.

The network layer MAY[109] support IP version 6 in accordance with IETF RFC 2460.

The network layer interface MUST[110] support differentiated service (Diffserv) code points in accordance with IETF RFC 3140.

Data Link Layer

The data link layer MUST[111] support Ethernet in accordance with IEEE 802.2/802.3 (Ethernet) and as shown in Figure 17.

The data link layer SHOULD[112] support Ethernet virtual LANs (IEEE 802.1Q).

The data link layer SHOULD[113] support IEEE 802.1D/Q.

The data link layer MUST[114] support PPP over Ethernet in accordance with IETF RFC 2516 and as shown in Figure 18

Logical Link Controller (LLC) Sublayer

The logical link controller sublayer subinterface MUST[115] support Ethernet in accordance with IEEE 802.2.

Medium Access Control (MAC) Sublayer

The medium access control sublayer subinterface MUST[116] support Ethernet in accordance with IEEE 802.3.

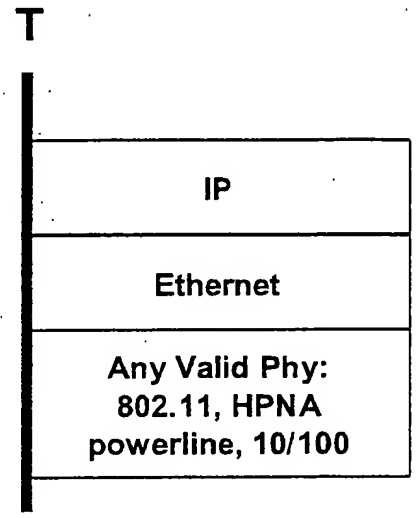

Figure 17 - IP over Ethernet

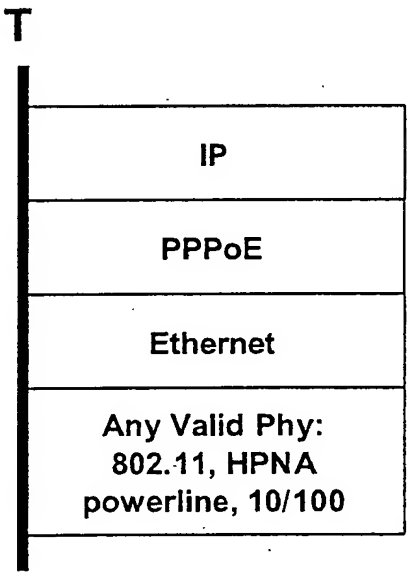

Figure 18 - IP over PPP over Ethernet

5 QUALITY OF SERVICE

5.1 Introduction

DSL architectures and products are predominately engineered for the support of best effort Internet traffic. Many NSPs desire the ability to improve their best effort product by using different levels of over subscription. Additionally, there are other market drivers pushing the Regional/Access Network to support differentiated services that require functionality beyond a best effort grade of service. Such services include telephony, video services, gaming, bandwidth on demand, and corporate VPN access as referenced in section 2.2. In order to support IP services effectively, the network MUST[117] be IP aware and provide support that scales as the number of DSL subscribers and the number of applications per subscriber increases.

5.1.1 Goals

The goal of this section is to describe the mechanisms for introducing:

- A method for providing different engineered performance to different networks – even for best effort traffic
- Per flow IP QoS into the Regional/Access Network

Both of these goals leverage the existing capital investments yet effectively meet the goals for supporting differentiated non-real time and real-time IP applications.

One goal of the architecture is to enable more flexible bandwidth allocation to customers. It is a goal to allow both the customer and the various Service Providers to participate in defining the bandwidth that will be made available to them via DSL. This bandwidth can be provided at different rates not only at provisioning time or via service orders, but also on demand in near real time using mechanisms like "turbo buttons" at NSP or ASP web interfaces, or by using signaling protocols. It should be noted that this is still best effort bandwidth – there is no guarantee that an application can make use of the maximum bandwidth, in other words there are no throughput guarantees – only that the possible maximum rate might be increased.

Real-time applications have concerns beyond bandwidth, like jitter and latency, which become harder to manage when the DSL line rate slows down. Other applications, while may not be real time, have delivery requirements (no packets dropped) that cannot be assured by bandwidth alone. It is a goal to manage multiple applications over a small number (1 or 2) of ATM PVC(s) between the DSL modem and BRAS and provide the characteristics that both real-time and non-real time applications require.

5.1.2 Assumptions

Existing Regional Networks have a large embedded base of ATM equipment that is not IP aware. This equipment will be leveraged to the extent that it is technically and economically feasible.

5.2 Traffic Engineering of Best Effort Service

Today's DSL access and Regional Networks are typically engineered to an over subscription ratio picked by the various providers. This has served the market well, but may need to be enhanced as service diversity expands and scope broadens. The concept for traffic engineering best effort service is that an NSP might be able to select an over subscription policy, and that the various NSPs can use that as a tool for providing different grades of service, even in an otherwise best effort model. Using this feature, one NSP may opt for highly over-subscribed infrastructure in order to provide an extremely cost-effective service, while a second NSP might choose a much less over subscribed approach in order to provide a better user experience or a premium service.

5.2.1 Theory of Operation

Traffic engineering (TE) makes use of MPLS TE, ATM VP or VC, and L2TP features in order to provide a specific over subscription rate for that NSP.

As shown in Figure 19, traffic flowing between NSP_1 and CPN_1 is shaped to a large asymmetric configuration through the Regional/Access Network. At the same time, traffic flowing between NSP_2 and CPN_2 is shaped to a smaller symmetric configuration. Finally, ATM or Diffserv techniques can be used at the A10-NSP interface in order to divide the total bandwidth at the interface among potentially disparate tunnel types that traverse it.

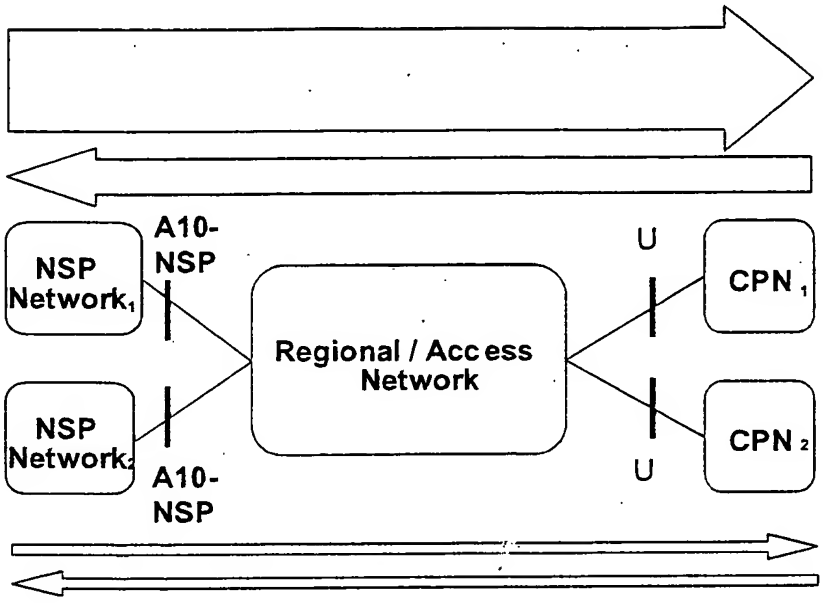

Figure 19 – Best Effort TE

5.3 QoS Architecture - A two-phased approach

While a signaled per flow IP QoS mechanism may ultimately be required for this architecture, the technical and economic feasibility of such a build out can not be justified in the near term. Instead, a 2-phase approach is suggested that leverages incremental IP awareness associated with ATM level traffic engineering. In the first phase, IP aware network elements are added to the network that in conjunction with ATM traffic engineering can manage IP flows through non-IP aware devices. The Diffserv model is leveraged to prioritize and shape traffic through ATM devices. The bandwidth that a subscriber receives will no longer be determined by the DSL synch rate alone. Instead, both the physical and IP layers will be leveraged. Most importantly, phase 1 significantly increases the IP layer functionality of the Regional/Access Network while not requiring massive re-deployment of capital and re-engineering of the network.

Phase 2, however, will require more enhancements to the Regional/Access Network by further increasing the IP capabilities. Policy-based IP QoS is introduced in this phase to allow mass customizability and per-application treatments.

5.3.1 Phase 1 QoS Mechanisms

Phase 1 largely leverages the existing broadband Regional/Access Network as shown in Figure 3. This network is generally IP unaware. In order to efficiently add IP awareness to the network without upgrading the ATM or Access node base, two enhancements are required: Within the network the BRAS is leveraged to provide IP aware handling of traffic and, similarly, at the customer premises new IP aware CPE capabilities are deployed.

One of the goals of this architecture is to provide differentiated services with IP QoS over a non-IP-aware layer 2 network. Since the layer 2 QoS features are not IP aware, they are left unused. Thus traffic from different IP QoS classes is put in the same queues in the layer 2 nodes. Since the layer 2 nodes cannot identify the different IP QoS types within a single queue, congestion $MUST_{[118]}$ be avoided in all layer 2 network elements at all times in order to retain IP QoS. Furthermore, IP QoS types that offer jitter management will also require that not only congestion is avoided in the L2 queues, but also that significant queuing delays are avoided as well. By avoiding congestion in the layer 2 network, its role is reduced largely to transport, and the switches are modeled like simple multiplexers. This means that the buffering mechanisms in the layer 2 nodes are avoided. Avoiding downstream congestion in the layer 2 network can be achieved by giving the BRAS full awareness of a logical

tree-based network topology. This topology is based on the actual physical and logical topology, but excludes resources that are used by other services (see section 5.3.1.2). The BRAS MUST$_{[119]}$ be aware of all potential congestion points in this constrained topology, as well as the trunk bandwidths and DSL synch rates. The BRAS MUST$_{[120]}$ make sure that no more traffic is inserted in the layer 2 network than is allowed according to its knowledge of the logical topology and customer policy constraints. This can be achieved using a hierarchical scheduling mechanism in conjunction with provisioning of services and policies in a way that remains aware of the topological network constraints.

The BRAS MUST$_{[121]}$ be able to police upstream both for traffic aggregates and for sub-classes of the aggregate using the same topology information that exists for the hierarchical scheduler. The BRAS SHOULD$_{[122]}$ support random differential drop behavior for upstream traffic aggregates and sub-aggregates based on class. Note that this is required because the RG just has a view of its own DSL line, and doesn't know about the DSL lines that belong to other RGs.

The expectation is that overall admission control for provisioning of bandwidth and the higher tiers of QoS will occur in a policy-based management system that will allow topology, access rates, and business service logic to be applied as part of the provisioning process. The BRAS and RG will enforce the resultant policies.

When a subscriber purchases a differentiated service, this service MUST$_{[123]}$ flow through the BRAS. To support differentiated services, the BRAS preserves IP QoS downstream through the access node and to the customer premises by means of packet classification, traffic shaping and hierarchical scheduling based on the logical tree-based network topology between the BRAS and the RG.

Once the BRAS is capable of managing the traffic flow through the access node, there is no need for access node to restrict a subscribers connection speed at layer one (ADSL synch rate). Instead, the access node should allow the ATU-R to synch up at its maximum rate. Access sessions will now be shaped and rate limited by the BRAS and can allow for multiple sessions to be individually shaped based on the subscribed service.

The BRAS MUST$_{[124]}$ support packet classification and scheduling in accordance with Diffserv.
The BRAS MUST$_{[125]}$ support hierarchical shaping, scheduling, and policing for the control of traffic through the access node and any other intervening devices that do not have IP awareness.
Implementations of hierarchical scheduling MUST$_{[126]}$ be resource efficient in the sense that any traffic MUST$_{[127]}$ be capable of using the subscriber bandwidth that has been allocated to that traffic class and that different classes should be able to make use of the unused subscriber bandwidth of other traffic classes.

The effectiveness of using hierarchical scheduling across non-IP aware devices decreases as the number of devices and the amount of non-BRAS controlled traffic increases. As a result, the BRAS function SHOULD$_{[128]}$ be located as close to the access node as possible from an ATM hop perspective. The daisy chaining SHOULD NOT$_{[129]}$ exceed a depth of more than two ATM switching/multiplexing points in the Access Node. Additionally, if the BRAS does not include ATM switching functions, then an additional layer of hierarchical scheduling MUST$_{[130]}$ be used to manage the trunk to the ATM switch.

The BRAS function MAY$_{[131]}$ be integrated into the access node, however one of the constraints of this architecture is that it must account for a large embedded base of access nodes that do not support this function.

In order to preserve an IP flow's characteristics, the customer CPE MUST$_{[132]}$ be involved in the QoS architecture. This is especially true when dealing with upstream traffic. This connection is typically the slowest link, and the most likely link to incur congestion and add delay and jitter within the service. To maintain fair but effective throughput over this link the RG MUST$_{[133]}$ support packet classification and scheduling in accordance with Diffserv. The RG MUST$_{[134]}$ also support a method of minimizing latency for EF traffic (e.g. fragmentation or MTU adjustment) that minimizes overhead, especially at times when no EF traffic is present.

The typical DSL customer is connected to the Regional/Access Network via a single ATM AAL5 PVC. This single PVC should be leveraged to the extent possible using the capabilities described above. Although the target architecture to support QoS enabled IP services seeks to utilize a single ATM AAL5 PVC per CPN, it is recognized that certain required network element features identified in this document have yet to be developed. In particular, dynamic packet fragmentation/MTU sizing in the CPE (needed to control jitter and delay for short packet/high priority applications) may trail the availability of other required network element features. In order to meet the demands of service descriptions previously identified in an acceptable timeframe, a second ATM PVC

MAY[135] be provisioned as an interim solution to provide a means to separate those application flows having tight jitter and latency requirements. This second PVC will require that DSL modems support multiple PVCs. For the service model proposed in this document, the number of PVCs per customer SHOULD NOT[136] exceed 2.

To support bandwidth on demand products or other differentiated services that implicitly require additional bandwidth on demand, a subscriber's access sessions MUST[137] be shaped and policed by the BRAS and RG instead of permitting cell insertion at the DSL line rate. This change is accompanied by changing the ATUs to allow them to synchronize at or near their maximum rate. Since this architecture allows for multiple simultaneous access sessions, it MUST[138] also be possible to independently modify the shapers and policers on each session. The policy data for the classification and shaping of traffic at the RG is provided at service configuration and is not a real time capability. The policy data for the classification and shaping of traffic at the BRAS can be provided at service configuration or may be dynamically configured.

Phase one assumes that the Regional/Access Network provider has established an IP-based architecture similar to that shown in Figure 4. This figure can be combined with Figure 2 in order to support the end-to-end view of the QoS-enabled network that follows. That combination is presented in Figure 20.

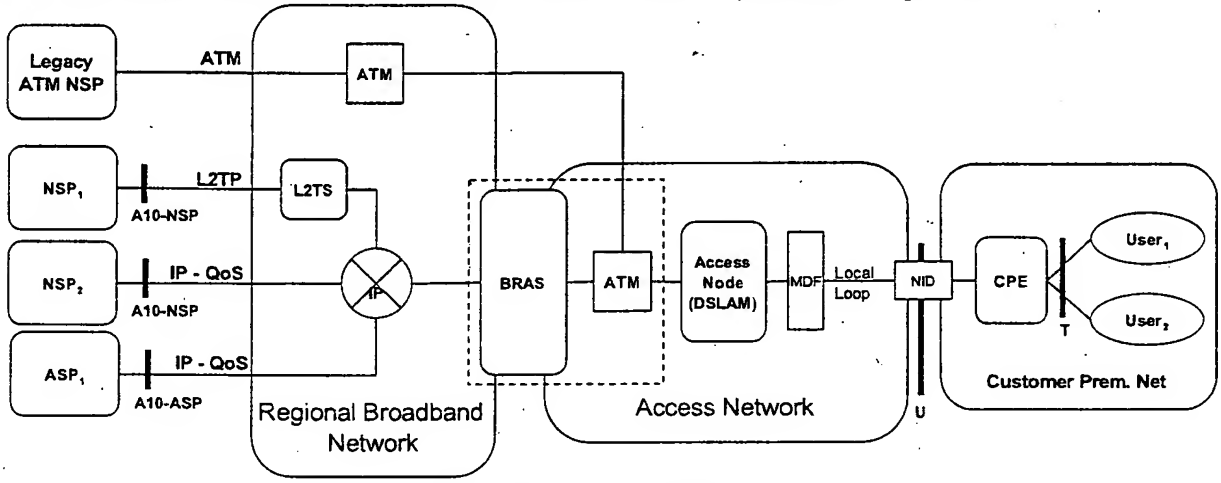

Figure 20 – QoS-enabled Network Topology

The two critical points where IP-QoS is managed are the BRAS and the CPE (RG). Intervening elements (like DSLAMs) are not envisioned to become layer-3 routers, and this architecture assumes that they will not be layer-3 aware when they manage congestion. This arrangement supports multiple business relationships and provides connectivity for various users to access various services without requiring all services to be provided by a single provider.

Phase 1 is characterized by Diffserv provided through static provisioning. Phase 2 describes a subsequent time with a dynamic mechanism for changing the Diffserv QoS parameters through the use of a policy-based networking enhancement.

Phase 1

Assumptions:

- In this phase there will be multiple BE NSP connections with few (1 or 2) EF sessions for real time applications (voice, Video conferencing). There will be little or no AF traffic – as applications would have to make use of static pre-configured AF classes rather than requesting one that suits the application.
- For phase 1 it is assumed that only one EF application per subscriber need be supported at a time (user performs the CAC across real-time applications). Within an application domain it is the application's responsibility to perform the CAC.
- Classification is performed at the RG on a session basis or accepted via markings attached to packets by the CPE

- Performing dynamic, per-application classification requires a 5-tuple classifier to be pushed down to the RG and is not likely in the short term.
- The DSLAM modems are allowed to sync near or at the max rate both upstream and downstream.
- Hierarchical scheduling is performed at the BRAS to provide IP QoS congestion mechanisms for the downstream path. Similar policing is performed in the upstream path.
- Packet-by-packet QoS requires being in the PTA (or bridged 2684) model at a given element.

Characteristics:

- Multi-user multi-destination is supported
- IP QoS is managed at the RG and BRAS
- The RG and BRAS are configured with common set of traffic profiles
- RG is configured by manufacturer or during installation (install CD)
- Statically assigned BE and EF queues will be supported in the RG.
    - Optional are statically assigned AF queues that could support 3 or 4 popular streaming arrangements or potential Gold/Silver/Bronze services. This option will require defining Diffserv classes that will be applicable across envisioned future services.
- Profile information defines the rate to which traffic should be shaped and the queuing behavior that should be used.
- Profile information will also determine the valid DSCPs.
- A small number of shaping profiles will be defined for the various connection speeds (e.g. 1.5x265; 1.5x 384; 384x384; 768x512)
- Sessions are individually shaped based on profile and share the aggregate DSL synch rate. If the total BW per profile exceeds the available sync rate then the traffic shares the BW in a "fair" manner among similar QoS service classes.
- If the RG initiated the session, and it is authenticated, then it is told which pre-provisioned profile to use. Various potential protocols and mechanisms to do this have been discussed at the DSLF. Note, if a CPE device behind the RG initiates a PPPoE session then it remains PPPoE through the RG, and is BE traffic by definition. (Even if it becomes a PTA connection at the BRAS)
- In either a PTA or L2TP model the BRAS will police traffic in the upstream direction and shape traffic in the downstream direction.
- BRAS shaping, policing, and marking is done on a per session basis, not per application. However, the diffserv queues can be arranged within an access session so that various aggregate service classes can be provided to applications that indicate which class of service they desire. The application needs to set the DSCP properly in order to make use of this function.
    - An end-to-end PPP session is given a uniform QoS treatment, but can be shaped (e.g. 1.5x256).
    - A single, additional PPP or 1483 session is used to access the ASP network.
- The BRAS profiles are updated through provisioning, not signaling, and may be indicated via RADIUS.
- New profiles are added/updated in the RG by the customer manually configuring the device or by downloading a new software image

5.3.2 Phase 2 QoS Mechanisms

As previously mentioned, Phase 2 is adds a dynamic mechanism for changing the Diffserv QoS parameters through the use of a policy-based networking.

Assumptions:

- Builds on the capabilities in phase 1.
- This phase enhances the granularity of the classification and population of policies in the BRAS and RG.
- Multiple sessions to multiple destinations, each with multiple applications that may require different QoS treatment

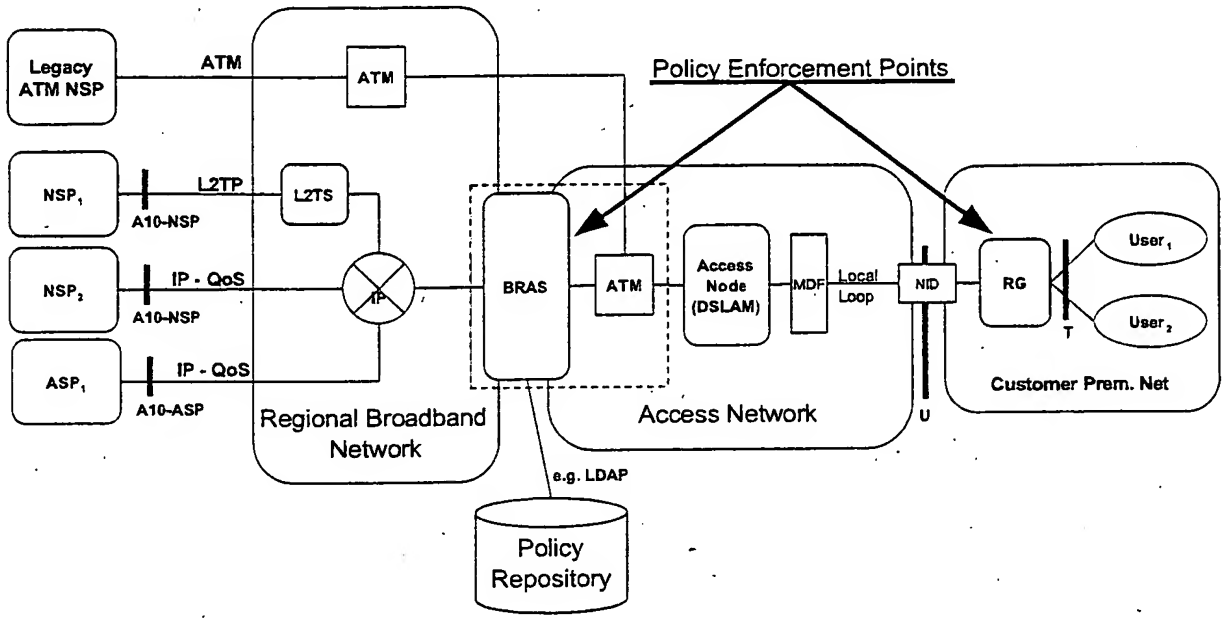

Figure 21 – Phase 2 with Policy-based profiles

Characteristics:

- When an NSP access session is authenticated the NSP MAY provide a profile indicator associated with that session in its response to the BRAS.
- Once the BRAS receives the profile indicator, it retrieves the full profile from the Policy Repository. Similarly, information is sent to the RG when it requests a profile. This step allows coordination among NSP and ASP profiles. Note also, that for some policy functions, the policy repository may be co-resident with the BRAS or RG.
- No single ASP authenticates the ASP access session, so a profile for that session is put together by the Policy Repository and is based on various ASP subscriptions associated with that access session.
- ASPs can update the profile either through subscription or through a dynamic protocol, like LDAP.
- Subscription profile information as well as DSL sync rate and user preferences are stored in the Policy Repository and accessed using a protocol, like LDAP.
- A policy manager is responsible for managing potentially conflicting ASP and NSP profiles and subscriptions and also creates billing data for services.
- The profile is populated in the elements in near-real time (no reset or "reboot" required).
- Diffserv marking and queuing behavior on RG is performed on 5-tuple matching (SA, DA, SP, DP, PI) as well as the mapping of existing marks and access sessions into various "equivalent" classes. For example, PPPoE access through a RG will continue to be given BE treatment.

5.3.2.1 Diffserv Requirements

RG

The RG requirements below only apply to the support of IP-QoS and should not be mistaken as a complete list of RG requirements needed in order to support this architecture.

The RG $SHOULD_{[139]}$ be the central point for controlling traffic within the customer premises and traffic destined for the Access Network.

The RG $MUST_{[140]}$ support Diffserv marking and remarking in accordance with IETF RFC 2474.

The RG MUST$_{[141]}$ support Diffserv queuing for the Assured Forwarding (AF) and Expedited Forwarding (EF) classes in accordance with IETF RFC 2597 and IETF RFC 3246 for carrying real time traffic. The exact AF classes supported and behaviors will be described in a future document.

The RG MUST$_{[142]}$ support multiple queues with the appropriate scheduling mechanism to effectively implement Diffserv queuing behaviors (e.g. strict priority, Weighted Fair Queuing).

The RG MUST$_{[143]}$ be configured with the classification parameters for mapping traffic into a given Diffserv Per Hop Behavior (PHB) during service configuration.

The RG MUST$_{[144]}$ support the capability to fragment AF and BE traffic in order to constrain the perturbing impact of AF and BE packets on EF traffic delay, for example using a mechanism such as MLPPP LFI [RFC1990].

The method of minimizing latency for EF traffic SHOULD$_{[145]}$ minimize overhead, especially at times when no EF traffic is present.

If multiple PVCs are provisioned at the ATU-R, the RG MUST$_{[146]}$ support the mapping between a Diffserv Code Point (DSCP) (low latency queue) and a specific PVC. (Using a PVC bundle is the desired way to meet this requirement.)

BRAS

The BRAS MUST$_{[147]}$ support Diffserv marking and remarking in accordance with IETF RFC 2474.

The BRAS MUST$_{[148]}$ be able to police the use of DSCPs received from customer traffic and remark traffic if it does not match the customer profile data – including potentially dropping unauthorized traffic.

The BRAS MUST$_{[149]}$ support Diffserv queuing for the Assured Forwarding (AF) and Expedited Forwarding (EF) classes in accordance with IETF RFC 2597 and IETF RFC 3246. The exact AF classes supported will be described in a future document. These queues are defined within the context of the DSLAM connectivity between the BRAS and the access node in affect managing the access node's downstream bandwidth.

The BRAS MUST$_{[150]}$ support multiple queues per user with the appropriate scheduling mechanism to effectively implement Diffserv queuing behaviors (e.g. strict priority, Weighted Fair Queuing).

The BRAS MUST$_{[151]}$ support the mapping of DSCP to MPLS LSP, VLAN, ATM VP, or other traffic engineering capabilities in the Regional Network.

The BRAS MUST$_{[152]}$ support the capability to fragment AF and BE traffic in order to constrain the perturbing impact of AF and BE packets on EF traffic delay, for example using a mechanism such as MLPPP LFI [RFC1990].

The method of minimizing latency for EF traffic SHOULD$_{[153]}$ minimize overhead, especially at times when no EF traffic is present.

If multiple PVCs are per subscriber are provisioned, the BRAS MUST$_{[154]}$ support the mapping between a Diffserv Code Point (DSCP) and a specific PVC. (Using a PVC bundle is the desired way to meet this requirement.)

5.3.2.2 Traffic Engineering Requirements

In order for the BRAS to effectively manage downstream IP traffic through layer 2 devices using the hierarchical scheduling model, the BRAS MUST$_{[155]}$ have awareness of all the traffic that is traversing those layer 2 elements. This can be accomplished in 2 ways. The first and most straightforward method is for all traffic destined for the access node to flow through the BRAS enabling it to manage the traffic accordingly. In this case the hierarchical scheduling model in the BRAS will be based on the full downstream trunk bandwidths and DSL synch rates. In cases where not all traffic flows through the BRAS, the resources that are not under the control of the BRAS MUST$_{[156]}$ be subtracted from the resources that the BRAS manages. The remainder of the resources on the trunks and DSL lines will be managed using the hierarchical scheduling model. The traffic that is not under the control of the BRAS MUST$_{[157]}$ be traffic engineered in a way that it cannot consume resources that the BRAS is controlling. Engineering around the BRAS incurs risk and must be done with care.

5.3.2.3 Admission Control

End-to-end QoS admission control is not required in this phase. Admission control for access network QoS (bandwidth on demand) is required. Application layer admission control will be predicated on service specific resources (such as availability of logical ports on servers and their knowledge of network topology). Furthermore, admission control may be provided in the provisioning aspect of a QoS policy.

6 SERVICE LEVEL MANAGEMENT

6.1 Introduction

Service Level Management is intended to provide 3 levels of benefit – increasing over time:

- To provide a list of the salient network performance and operational metrics that might be used in a Service Level Objective (SLO) or Service Level Agreement (SLA).
- To provide a standard definition of such metrics so that its meaning would be common when used by various providers.
- To provide extreme values that are driven by architectural considerations where applicable. For example, while it is NOT the intention of this document to set the SLO or SLA for Network Delay (Latency), any network that purports to support Voice over IP (VoIP) will need to have a maximum delay that is within the bounds necessary to support VoIP.

6.2 Network Performance Metrics

1. **Network Availability** - The percent of time that the Regional/Access Network is available for subscribers to connect. This metric is defined on some time basis, such as a month, a week, or a year. An SLA should also specify not the entire network but the section of the network for which the Regional/Access Network Provider is responsible. For example, the Regional/Access Network Provider is not responsible for NSP problems.
2. **Network Delay (Latency)** – The time it takes for a data packet to traverse the Regional/Access Network, from end-to-end or edge-to-edge. Latency is defined in milliseconds and can be a one-way or round-trip delay.
3. **Message Delivery** - The ability of the Regional/Access Network to transmit traffic at the negotiated speed. Some applicable measurements are packet loss). These metrics must have a time base as well.
4. **Network Jitter** – The variance of network latency. Jitter is defined in milliseconds.

6.3 Operational Metrics

1. **Mean Response Time** - The time it takes the Regional/Access Network Provider to respond to submitted reports of trouble
2. **Mean Time to Restore Service** – The measurement of the Regional/Access Network Provider's ability to restore service within the negotiated interval
3. **Ordering System Reliability** – The measurement of the consistent availability of ordering system.
4. **End-User Installation Guarantee** – The measurement of the Regional/Access Network Provider's ability to meet negotiated order due dates.

7 SERVICE MANAGEMENT

The architecture proposed in this document clearly needs management systems to provide the controls necessary to support the underlying service "building blocks". The following lists are examples of new data points that management systems $MUST_{[158]}$ support. Network elements and Service Providers will use these new data elements for service provisioning and data delivery. It is expected that the Operations and Network Management working group of the DSL Forum will provide contributions to augment this section.

7.1 Subscribers

Because of the changes in how DSL is provisioned and managed, there are a number of new data points that MUST[159] be tracked for each subscriber. Among these are:

- Maximum sustainable subscriber bandwidth
- Maximum number of sessions allowed
- Permitted destinations
- Default protocol
- Default destination
- Default bandwidth
- Single host or subnet needed
- Restricted subscriber (single destination only)
- Total reserved bandwidth

7.2 Service Providers

Because of the changes in how DSL is provisioned and managed, there are more details needed per Service Provider. When various choices listed for an option, these are to be considered as examples only and not a definitive list of the choices for a given option.

- Minimum bandwidth needed
- Minimum QoS level
- Various protocol metrics
- Subscriber protocol (IP, PPPoE)
- Protocol (IP, L2TP, ATM)
- Authentication
- IP address assignment
- Transport
- Maximum simultaneous sessions

GLOSSARY

AAA	Authentication, Authorization, and Accounting
AAL5	ATM Adaptation Layer 5
ADSL	Asymmetric Digital Subscriber Line
AF	Assured Forwarding
API	Application Program Interface
ARP	Address Resolution Protocol
ASP	Application Service Provider
ATM	Asynchronous Transfer Mode
ATMARP	ATM Address Resolution Protocol
ATMF	ATM Forum
ATU-C	Access Termination Unit - Central Office (at Access Network end)
ATU-R	Access Termination Unit - Remote (at customer end)
B-NT	Broadband Network Termination
BE	Best Effort
BGP	Border Gateway Protocol
BoD	Bandwidth on Demand
BRAS	Broadband Remote Access Server
CBR	Constant Bit Rate
CO	Central Office
COPS	Common Open Policy Service
CoS	Class of Service
CPE	Customer Premises Equipment
CPN	Customer Premises Network
CSP	Corporate Service Provider
DHCP	Dynamic Host Configuration Protocol
Diffserv	Differentiate Services
DLC	Digital Loop Carrier
DNS	Domain Name Service
DS1	Digital Signal level 1 (1.544 Mbps)
DSCP	Differentiated Services (Diffserv) Code Point
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
EF	Expedited Forwarding
ESP	Encapsulating Security Payload
FQDN	Fully Qualified Domain Name
GFR	Guaranteed Frame Rate
iBGP	internal Border Gateway Protocol
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IKE	Internet Key Exchange
IP	Internet Protocol
IPsec	Secure Internet Protocol
ISP	Internet Service Provider
ITU-T	International Telecommunications Union - Technical
L2TP	Layer 2 Tunneling Protocol
L2TS	Layer 2 Tunnel Switch
L2oMPLS	Layer 2 over MPLS
LAC	Layer 2 Access Concentrator
LAN	Local Area Network
LD	Long Distance
LDAP	Lightweight Directory Access Protocol
LER	Label Edge Router
LLC	Logical Link Control

LSP	Label Switched Path
LNS	L2TP Network Server
MAC	Medium Access Control
MARS	Multicast Address Resolution Server
MASS	Multi-Application Selection Service
MBGP	Multicast Boarder Gateway Protocol
MPEG	Motion Pictures Expert Group
MPLS	Multi-Protocol Label Switching
MS/MD	Multi Session / Multi Destination Service
MTU	Message Transfer Unit
NAPT	Network Address Port Translation
NG-DLC	Next Generation Digital Loop Carrier
NHRP	Next Hop Resolution Protocol
NSP	Network Service Provider
OC3	Optical Carrier 3
OSPF	Open Shortest Path First
PC	Personal Computer
PHB	Per Hop Behavior
PHY	Physical Layer
POP	Point of Presence
POS	Packet over SONET
PPP	Point-to-Point Protocol
PPPoA	Point-to-point Protocol over ATM
PPPoE	Point-to-Point Protocol over Ethernet
PTA	PPP Terminated Aggregation
PVC	Permanent Virtual Circuit
PVP	Permanent Virtual Path
QoS	Quality of Service
RADIUS	Remote Access Dial-In User Service
RAM	Remote Access Multiplexer
RFC	Request For Comments
RG	Routing Gateway
RRP	Resource Request Protocol
RSVP	ReSource reserVation Protocol
RT-DSLAM	Remote Digital Subscriber Line Access Multiplexer
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SLO	Service Level Objective
SNAG	Service Network Architecture Group (DSL Forum)
SONET	Synchronous Optical Network
SVC	Switched Virtual Circuit
TCP	Transmission Control Protocol
TE	Traffic Engineering
TR	Technical Report (DSL Forum)
TV	Television
UBR	Unspecified Bit Rate
UDP	User Datagram Protocol
VBR-nrt	Variable Bit Rate - non-Real Time
VBR-rt	Variable Bit Rate - Real Time
VC	Virtual Circuit
VCC	Virtual Circuit Connection
VLAN	Virtual Local Area Network
VoD	Video on Demand
VP	Virtual Path
VPC	Virtual Path Connection
VPN	Virtual Private Network

VoBB	Voice over Broadband
VoIP	Voice over Internet Protocol
WFQ	Weighted Fair Queuing

APPENDIX A REFERENCES

[1] DSL Forum TR-010, "Requirements & Reference Models for ADSL Access Networks: The "SNAG" Document"

[2] DSL Forum TR-025, "Core Network Architecture for Access to Legacy Data Networks over ADSL"

[3] DSL Forum TR-032, "CPE Architecture Recommendations for Access to Legacy Data Networks"

[4] DSL Forum TR-037, "Auto-Configuration for the Connection Between the DSL Broadband Network Termination (B-NT) and the Network using ATM"

[5] DSL Forum TR-042, "ATM Transport over ADSL Recommendation (Update to TR-017)"

[6] DSL Forum TR-043, "Protocols at the U Interface for Accessing Data Networks using ATM/DSL"

[7] M. Kaycee, G. Gross, A. Lin, A. Malis, J. Stephens, "PPP over AAL5," IETF RFC 2364, July 1998

[8] Skwoler, et. al, "The PPP Multilink Protocol (MP)," IETF RFC 1990, August 1996

[9] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", IETF RFC 2474, December 1998.

[10] L. Mamakos, et al, "A Method for Transmitting PPP Over Ethernet", IETF RFC 2516, February 1999

[11] D. Grossman, J. Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5", IETF RFC 2684, September 1999

[12] W. Townsley, et al, "Layer Two Tunnelling Protocol (L2TP)" IETF RFC 2661, August 1999

[13] J. Heinanen, R. Guerin, "A Single Rate Three Color Marker" IETF RFC 2697, September 1999

[14] J. Postel, J.K. Reynolds, "Standard for the transmission of IP datagrams over IEEE 802 networks", IETF RFC 1042, February 01, 1988.

[15] S.E. Deering, "Host extensions for IP multicasting", IETF RFC 1112, August 01, 1989.

[16] W. Fenner, "Internet Group Management Protocol, Version 2", IETF RFC 2236, November 1997.

[17] T. Bates, Y. Rekhter, R. Chandra, D. Katz, "Multiprotocol Extensions for BGP-4", IETF RFC 2858, June 2000.

[18] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", IETF RFC 2474, December 1998.

[19] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski., "Assured Forwarding PHB Group", IETF RFC 2597, June 1999.

[20] D. Black, S. Brim, B. Carpenter, F. Le Faucheur, "Per Hop Behavior Identification Codes", IETF RFC 3140, June 2001.

[21] B. Davie, A. Charny, J.C.R. Bennet, K. Benson, J.Y. Le Boudec, W. Courtney, S. Davari, V. Firoiu, D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", IETF RFC 3246, March 2002.

[22] D. Grossman, "New Terminology and Clarifications for Diffserv", IETF RFC 3260, April 2002.

[23] The ATM Forum "Traffic Management Specification Version 4.1", AF-TM-0121.000, March 1999.

APPENDIX B Informative Example of Queuing Architectures for RG and BRAS

B.1 Example Queuing Architecture for RG

The queuing and scheduling discipline envisioned upstream for the RG is shown in Figure 24.

There are multiple access sessions supported in this model, however, all traffic is classified and scheduled in a monolithic system. So, while it might appear at first that the Diffserv queuing and scheduling might apply only to IP-aware access – in fact all access, IP, Ethernet, or PPP is managed by the same system that adheres to the Diffserv model.

For example, at the bottom of the figure, BE treatment is given to the non-IP-aware access sessions (PPPoE started behind the RG or delivered to an L2TP tunnel delivery model). This queue might be repeated several times in order to support fairness among multiple PPPoE accesses – or it may be a monolithic queue with separate rate limiters applied to the various access sessions.

The PTA access is a single block of queues. This is done because NSP access typically works with a single default route to the NSP, and managing more than one simultaneously at the RG would be perilous. The $\sum$ rate limiter would limit the overall access traffic for a service provider.

Rate limiters are also shown within the EF and AF service classes because the definition of those Diffserv types is based on treating the traffic differently when it falls into various rates.

Finally, at the top of the diagram is the ASP access block of queues. In phase 1A, these queues are provisioned and provide aggregate treatment of traffic mapped to them. In phase 1B, it will become possible to assign AF queues to applications to give them specific treatment instead of aggregate treatment. The EF service class may also require a high degree of coordination among the applications that make use of it so that it's maximum value is not exceeded.

Notable in this architecture is that all the outputs of the EF, AF, and BE queues are sent to a scheduler (S) that pulls traffic from them in a strict priority fashion. In this configuration EF traffic is, obviously, given highest precedence and BE is given the lowest. The AF service classes fall in-between.

Note that there is significant interest in being able to provide a service arrangement that would allow general Internet access to have priority over other (bulk rate) services.[1] Such an arrangement would be accomplished by assigning the bulk rate service class to BE and by assigning the default service class (Internet access) as AF with little or no committed information rate.

Given this arrangement, the precedence of traffic shown in the figure is arranged as:

1. EF – red dotted line
2. AF – blue dashed line (with various precedence among AF classes as described in RFC2597)
3. BE – black solid line

[1] This "bulk rate" service class would typically be used for background downloads and potentially for peer-to-peer applications as an alternative to blocking them entirely.

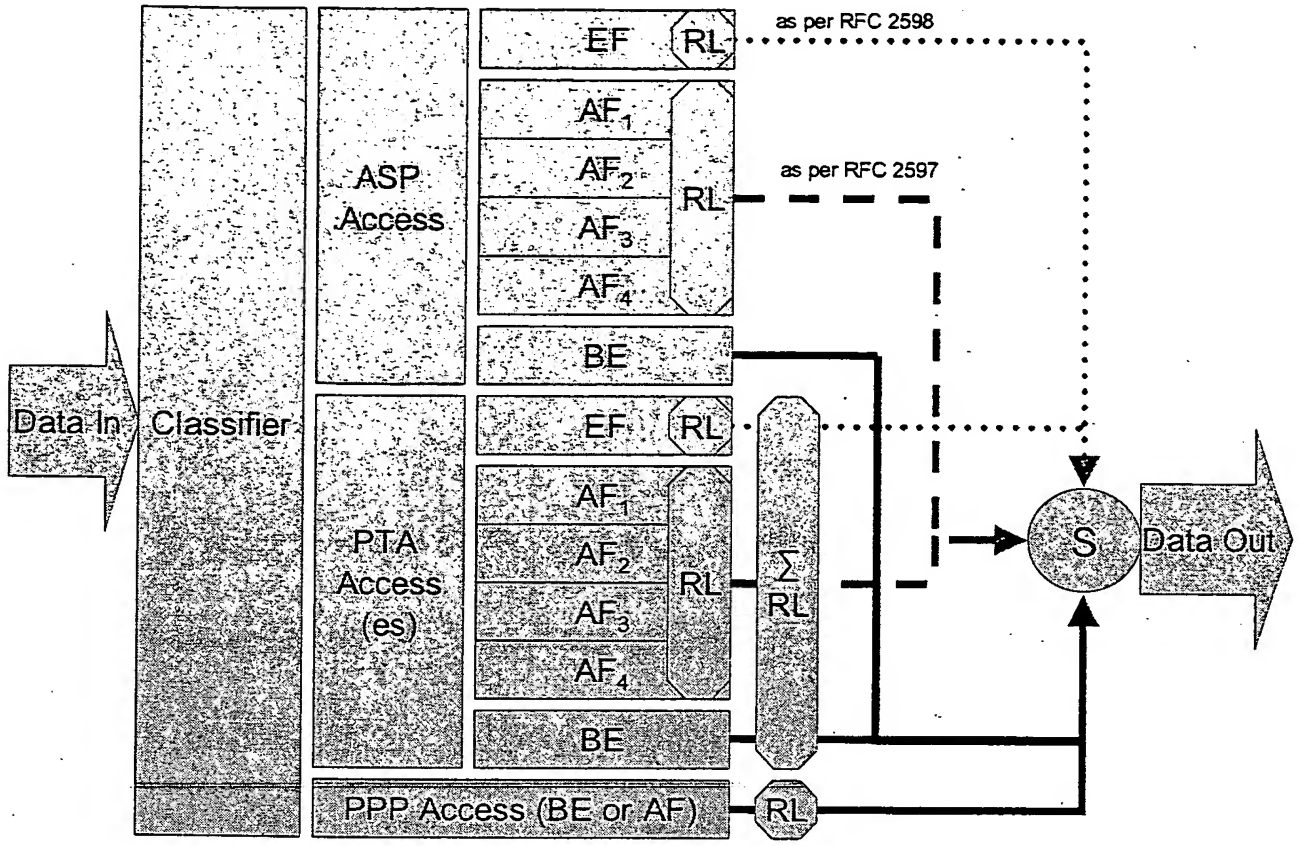

Figure 22 – Queuing and Scheduling Example for RG

In Figure 22 the following abbreviations apply:

ASP – Application Service Provider
PTA – PPP Terminated Aggregation
PPP – Point-to-Point Protocol
EF – Expedited Forwarding – as defined in RFC 3246
AF – Assured Forwarding – as defined in RFC 2597
BE – Best Effort forwarding
RL – Rate Limiter
∑RL – Summing Rate Limiter (limits multiple flows)
S – Scheduler

B.2 Example Queuing Architecture for a BRAS that can also switch ATM

An example of a queuing and scheduling discipline for a BRAS that meets the hierarchical shaping/scheduling requirements envisioned downstream is shown in Figure 24. Note that in this example, the BRAS is also an ATM switch, although the ATM switching capability is not essential for all BRAS designs.

There are multiple access sessions supported in this model, however, all traffic is classified and scheduled in a monolithic system. So, while it might appear at first that the Diffserv queuing and scheduling might apply only to IP-aware access – in fact all access, IP, Ethernet, PPP, and even ATM is managed by the same system that adheres to a combination of queuing disciplines taken from ATM and the Diffserv model. Note that the ATM disciplines are for backward compatibility, and don't otherwise interact with the Diffserv disciplines.

The BRAS will need to provide a congestion management function that will allow the synthesis of IP QoS through downstream elements that are not QoS aware. Accomplishing this is envisioned as a marriage of IP and ATM technologies with ATM and WFQ scheduling performed against diffserv and ATM queues. At a very high level, the queuing architecture desired for the BRAS can be described as IP DiffServ classification and queues mated to a slightly enhanced ATM scheduler. This results in emitting (shaping) ATM cells into the downstream network according to their VC contracts and ATM traffic engineering requirements, and so that no congestion occurs on the downstream links and systems. The result is that congestion queues in the BRAS, and eventual data discard occurs in packets being dropped from the DiffServ queues according to their precedence.

Figure 23 is provided as a reference to reinforce the problem and to provide exemplary infrastructure to show how the queuing system works.

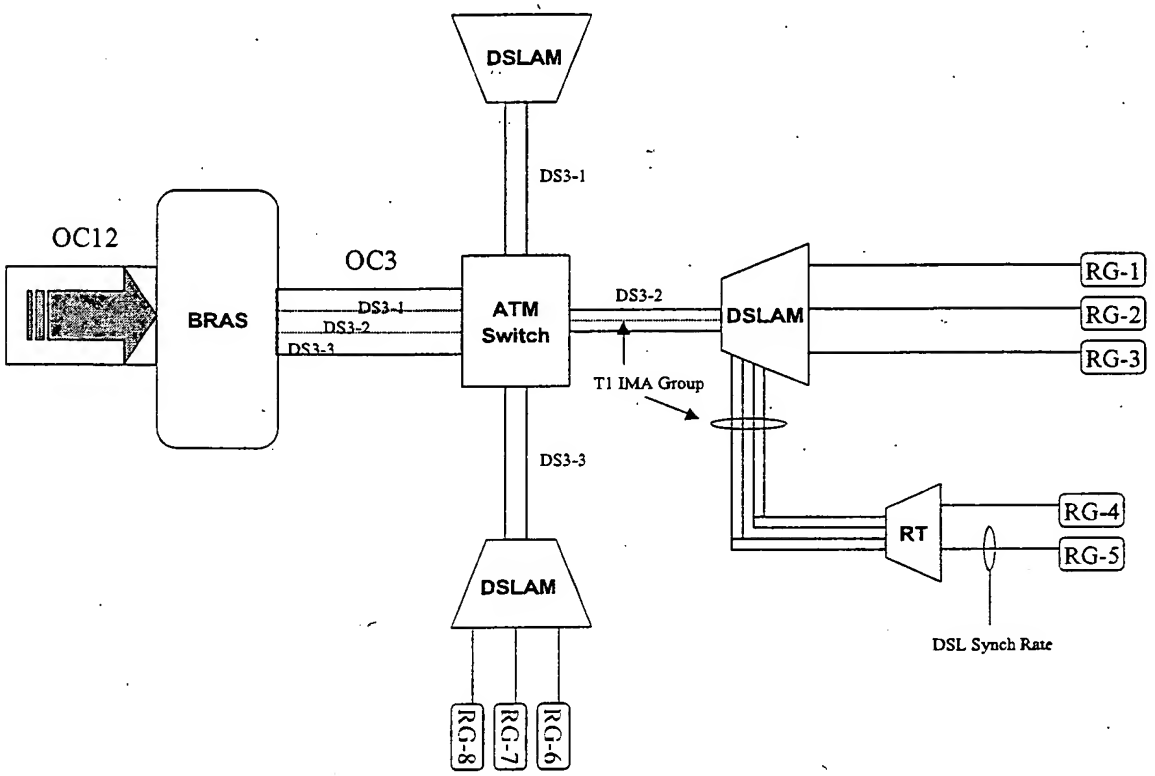

Figure 23 – Reference Topology for Queuing and Scheduling Example for a BRAS that can also switch ATM

In this example the BRAS is on the left and uses a central ATM switch to multiplex access to 3 DSLAMs, at the top, right, and bottom. The DSLAM on the right has an additional RT unit daisy-chained behind it using a T1 IMA group. Various RGs are behind the ADSL lines at differing sync rates. As stated in WT81, there is an assumption that congestion in this network never occurs in the fabric of the ATM switch, DSLAMs, or RT units, and always occurs through the over-subscription of transport links. In this example, those links would be:

1) OC 3 between BRAS and ATM switch
2) DS3 between ATM switch and DSLAMs
3) T1 IMA between DSLAM and RT
4) DSL loop to the RGs

Now, we observe traffic entering the BRAS and its queuing discipline, and see the following:

1) First, traffic is classified in a similar way to what was described for the RG. One notable exception being legacy ATM traffic, which is queued according to the class associated with the VC.

2) It is then policed, or rate limited, according to the services associated with the queues (if any). Again with an ATM exception of applying ATM-appropriate disciplines, such as CBR, VBR or UBR.

Traffic remains in the queues until it is scheduled for delivery. If congestion would occur in the BRAS or on a downstream link, then the queues for that traffic fill according to their discipline.

1) The scheduler is best described in reverse. First, the egress port of the BRAS is scheduled to the port rate (OC3 in this example). At this level, the scheduler is set for a WFQ algorithm, weighted according to the data rates of the VPs that are scheduled. Traffic is "pulled" from the subordinate schedulers in priority (as described for the RG scheduler) but with the limitations set by the various subordinate schedulers.
2) Then each ATM VP is scheduled. In this case there are 3 DS3 VPs that each lead to a different DSLAM and are scheduled to the DS3 rate. The schedulers are set to work in a similar way to the egress port scheduler.
3) In a departure from a typical ATM device, an additional layer of hierarchy is defined for "groups" of VCs in order to account for bandwidth constraints beyond the DSLAM. This can occur with DLC-based and RT-based DSLAMs that typically us IMA groups daisy-chained into Co-based DSLAMs. In this example, the VC Group Scheduler accounts for the T1 IMA group to the RT.
4) The next stage is the scheduler for the ATM VC. This scheduler works almost exactly like the RG. In the (optional) case where 2 PVCs are used the bandwidth of the DSL line is divided between the 2 PVCs instead of being directly assigned.
5) Finally, the queues within a given access session are scheduled to a maximum rate assigned to the access session. Initially static, the limit eventually becomes profile-driven through the policy repository.

As was described for the RG queuing architecture, all the outputs of the EF, AF, and BE queues are sent to a (hierarchical) scheduler (S) that pulls traffic from them in a strict priority fashion. Similar to the description of the RG queuing, a configuration may create the opportunity to establish access types with a lower priority than existing Internet access.

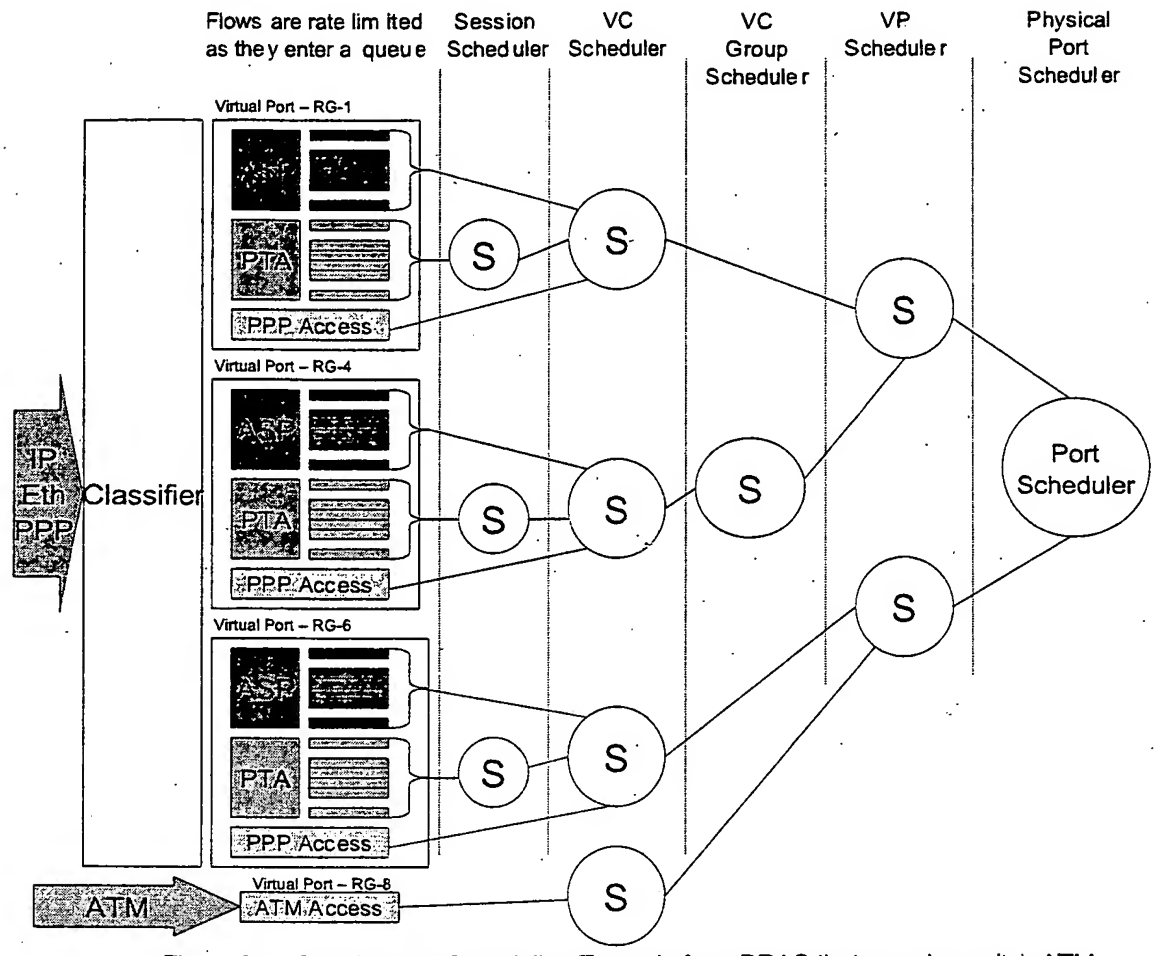

Figure 24 – Queuing and Scheduling Example for a BRAS that can also switch ATM

In Figure 24 the following abbreviations apply:

ASP – Application Service Provider
ATM – Asynchronous Transfer Mode
PTA – PPP Terminated Aggregation
PPP – Point-to-Point Protocol
S – Scheduler

APPENDIX C Informative Appendix on Signaled QoS

This appendix captures the concepts and planning for a potential follow-on signaled QoS mechanism. While it is uncertain that this phase will be required, concepts are brought forward to provide a perspective of how it would interact with the QoS mechanisms defined in this specification. The exact signaling protocol remain an item of debate, so this section refers to it with the abstract term Resource Reservation Protocol (RRP) and collects attributes that are likely to become requirements when the protocol becomes defined.

C.1 Signaled QoS Mechanisms

The architecture for additional QoS enhancements is introduced in this section. This section is included for illustrative purposes and may be further defined in future documents.

Signaled QoS adds per IP flow resource reservation capabilities in the Regional/Access Network. This step continues to leverage the RG and BRAS as the IP QoS managers of the access network. Rather than simply managing the aggregate scheduling of Diffserv resources, the BRAS will be able to perform per flow admission control ensuring that resources are never over-booked. Diffserv aggregate traffic treatments may continue to be used beyond the BRAS toward the regional network for scalability reasons. Keeping per flow resource reservation limited to the access portion of the Regional/Access Network could limit scalability/performance issues known with prior end-to-end reservation schemes.

In this phase:

- Applications, located in any of the reference networks, request service or resources of the Regional/Access network (e.g. through RRP).
- The RG and BRAS are involved in requests for services and resources in the network based on a per-application need (e.g. they monitor or proxy RRP messages).
- The BRAS acts like an RSVP border proxy and queries the policy repository. It responds based on the network availability of traffic engineered resources (MPLS – TE, ATM VP, etc) and customer profile.
- QoS service profiles can be applied to the BRAS and RG based on the requested application need.

C.1.1 Signaled QoS Assumptions

BRAS

The BRAS may support a RRP for the assignment of resources. When resources are not available at any point under its control the BRAS would reject the request and provide feedback to the initiating host.

The BRAS would need to know the DSL sync rates of the ATU-Rs that are connected to the access nodes that it manages. Based on a given ATU-R's DSL synch rate and customer profile the BRAS would manage the admission of sessions to that customer premises. An external policy/management server could feed this information to the BRAS.

The BRAS might intercept RRP and other application layer (e.g. SIP) messages that are not addressed to it and use these messages in making admission decisions.

The BRAS would support mapping reservation requests into Diffserv PHBs and managing the PHBs as reservable resources.

CPE

The CPE assumptions below only apply to the support of differentiated services.

The CPE requesting differentiated services could be integrated with the ATU-R. Non-integrated CPE devices will also be supported (e.g. IP Phones, PC running video conferencing software, set top boxes, etc).

The CPE would need to support IP layer signaled QoS via a RRP. These messages would be addressed to the destination host and not to the BRAS.

The CPE would not make any admission decisions.

C.1.2 Diffserv Assumptions

BRAS

The BRAS will accept policy information regarding how to manage Diffserv signaled flows from an external entity.

CPE

If the signaling messages indicate the DSCP to be used by a session requesting access, the CPE would then use the specified DSCP.

The CPE will also accept policy information regarding how to manage Diffserv signaled flows from an external entity.

C.1.3 Traffic Engineering Requirements

The RRP mechanism described only has resource knowledge of the local access network and does not have an end-to-end picture of the connection. As a result, the interconnection network within the Regional/Access Network (beyond the BRAS) would be engineered to provide support for enhanced services in aggregates. It is expected that within the core of the Regional/Access Network that aggregate traffic engineering techniques can efficiently serve the needs of enhanced applications.

C.1.4 Admission Control

Per-flow admission control is envisioned at the BRAS. Admission decisions are made based on resource availability AND subscriber profile data. Both of these parameters could be sent to the BRAS via an external policy/provisioning server.

Application level admission control can also be applied in addition to the network based admission control.